

W P 1231

**L'impact des politiques d'exonérations territoriales :
méthodes d'évaluation et résultats**

Nathalie Havet

October 2012

GATE Groupe d'Analyse et de Théorie Économique Lyon-St Étienne

93, chemin des Mouilles 69130 Ecully – France

Tel. +33 (0)4 72 86 60 60

Fax +33 (0)4 72 86 60 90

6, rue Basse des Rives 42023 Saint-Etienne cedex 02 – France

Tel. +33 (0)4 77 42 19 60

Fax. +33 (0)4 77 42 19 50

Messagerie électronique / Email : gate@gate.cnrs.fr

Téléchargement / Download : <http://www.gate.cnrs.fr> – Publications / Working Papers

L'impact des politiques d'exonérations territoriales : méthodes d'évaluation et résultats

Nathalie Havet *
GATE Lyon-Saint Etienne
Université de Lyon, CNRS

Résumé

Cet article propose une revue de la littérature sur l'efficacité des expériences de politiques d'exonérations ou d'allègements de charges sociales ciblées sur les territoires aux Etats-Unis et en France. Elle montre que leurs résultats sont mitigés en termes de créations nettes d'emplois, de créations d'entreprises et de retours à l'emploi des chômeurs de ces zones. Pour la France, les évaluations sur les Zones Franches Urbaines (ZFU) concluent à un faible impact positif sur l'activité locale alors que celle sur les Zones de Redynamisation Rurale (ZRR) à aucune retombée significative. Le régime dérogatoire dans les DOM, certes le plus coûteux, n'a malheureusement pas encore fait l'objet d'évaluation ex-post faute de données fiables disponibles. Les travaux existants montrent par ailleurs que l'évaluation de tels dispositifs doit relever deux défis pour être de bonne qualité : d'une part, définir convenablement un groupe de contrôle auquel les zones bénéficiant des dispositifs pourront être comparées ; d'autre part, rendre compte correctement, dans les données analysées, du périmètre des zones traitées.

Abstract : The impact of the place-based tax exemptions : evaluation strategies and results. This paper provides a survey on the efficiency of locally-targeted tax incentives in revitalizing distressed areas in the U.S. and in France. It shows mixed results on labour market performances (net job creations, business creations, ability of unemployed workers in these areas to find a job). In France, the evaluations of the *Zones Franches Urbaines* (ZFU) program find a small positive effect on local economic activity while the evaluations of the *Zones de Redynamisation Rurale* (ZRR) find no significant impact. Unfortunately, the costly special tax exemptions program for the overseas French departments still waits for its ex-post evaluation because of a lack of reliable data. The literature also shows that each evaluation of the place-based tax exemptions programs has to take up two main challenges : constructing a good control group for the targeted areas to obtain relevant comparisons and using sufficiently disaggregated geographic data to precisely define the boundaries of targeted and no-targeted zones.

Mots clés : évaluation de politiques publiques, méthode du score de propension, méthodes par appariement, zone franche sensible (ZUS), exonérations

Keywords : public policy evaluation, matching, propensity score, differences-in-differences, enterprise zones, tax exemptions

JEL : J08, J48.

*Université de Lyon, Lyon, F-69007, France ; CNRS, GATE Lyon St Etienne, Ecully, F-69130, France, Université Lyon 2, Lyon, F-69007, France. GATE : 93, Chemin des Mouilles, B.P. 167, 69131 Ecully, France. Tél : 04 72 86 60 83. Fax : 04 72 86 60 90. E-mail : havet@gate.cnrs.fr

1 Introduction

Les politiques de discrimination positive mettent en œuvre des actions volontaristes préférentielles afin de corriger des situations d'inégalités avérées et persistantes. Nées aux Etats-Unis au milieu des années 1960 suite au mouvement des droits civiques, elles se sont, à leurs débuts, principalement attachées dans ce pays à améliorer l'accès à l'emploi et à l'éducation supérieure des minorités ethniques et des femmes. Toutefois, à partir des années 1980 aux Etats-Unis et en Europe, des politiques préférentielles, non fondées sur l'appartenance raciale ou ethnique d'une minorité ou à un genre, mais sur des critères socioéconomiques ont vu le jour. Elles ont été motivées par la reconnaissance d'inégalités persistantes liées à l'appartenance socioéconomique telles que la ségrégation urbaine, la ségrégation scolaire, les inégalités devant l'emploi ou l'accès au logement [Maguain (2006)]. Dans cette logique, la France a par exemple mis en place au début des années 1980 des *zones d'éducation prioritaire* (ZEP), destinées à allouer des moyens supplémentaires aux établissements scolaires de certains quartiers défavorisés à faibles réussites scolaires. Or, devant le constat que le territoire était un reflet des disparités socioéconomiques, des mesures de discrimination positive territoriale ont été instaurées. En France, depuis le début des années 1990, et la mise en place des *zones urbaines sensibles* (ZUS), certains quartiers qui cumulent de nombreux handicaps (chômage de masse, précarité, habitat dégradé, faible accès aux services publics) bénéficient d'une réglementation plus favorable (ex : avantages fiscaux) que d'autres pour les entreprises qui décident de s'y implanter. Ces dispositifs ciblés sur des territoires particuliers ont d'une part trouvé leur justification dans des travaux d'économie géographique consacrés au *spatial mismatch*, montrant qu'il est dommageable pour les travailleurs d'être physiquement éloignés des lieux d'emplois [Gaschet & Gaussier (2003), Gobillon, Selod & Zenou (2007), Duguet, L'Horty & Sari (2009)]. En effet, cette situation accroît considérablement leurs coûts de recherche d'emploi et leurs coûts de transport domicile-travail et nuit par là-même à leurs chances de décrocher un emploi stable. D'autre part, il a été observé que la qualité des réseaux sociaux dans ces zones sensibles était vraiment mauvaise et donc que les personnes qui y étaient au chômage ne pouvaient en général pas s'appuyer sur leurs connaissances pour les aider à trouver un emploi [Korsu & Wenglenski (2010)].

Cette discrimination positive territoriale n'a ensuite cessé de prendre de l'ampleur en France. Au croisement de l'aménagement du territoire, de la politique de la ville, de l'accompagnement des mutations économiques et du soutien à l'Outre-mer, plusieurs dispositifs zonés géographiquement se sont accumulés pour répondre aux difficultés spécifiques de certains territoires. Ils partageaient tous le même objectif : promouvoir le développement économique local de ces zones en y favorisant l'implantation d'entreprises et l'embauche des salariés y résidant. C'est pourquoi, la plupart des mesures de discrimination positive territoriale ont pris la forme de dispositifs d'exonérations fiscales et sociales, ciblés sur certaines zones géographiques bien précises. Au principe d'égalité des résultats (notamment dans l'accès à l'emploi) propre aux politiques de discrimination positive, elles ont combiné l'idée ancienne que la réduction des prélèvements fiscaux et sociaux était la clé pour encourager l'emploi. D'ailleurs, depuis 1993,

la réduction générale des cotisations patronales de sécurité sociale pour les bas salaires constitue une orientation fondamentale de la politique de l'emploi en France qui n'a pas été remise en cause depuis, même si certaines de ses modalités (salariés bénéficiaires, taux d'exonérations) ont été adaptées au fil des années. Des allègements de charges spécifiques ont même été ciblés soit sur certaines catégories de population tels que les jeunes (apprentissage, contrats aidés, *etc*) ou les chômeurs de longue durée (contrats initiative-emploi, emploi-solidarité emploi-consolidé, aides à la création ou à la reprise d'entreprises), soit sur certains types d'emplois (travail à temps partiel) ou sur certains secteurs d'activité (hôtel-café-restaurant, services à la personne). Les exonérations territoriales sont emblématiques dans cette logique de réduction du coût du travail : les territoires présentant des problèmes d'emplois ou d'attractivité importants bénéficient d'allègements de charges particuliers.

La discrimination positive territoriale est notamment montée en charge en 1996 avec la mise en œuvre du "Pacte de relance pour la ville", qui outre l'extension de la liste des ZUS, a créé des *zones de redynamisation urbaine* (ZRU) et des *zones franches urbaines* (ZFU), accompagnées d'un dispositif d'allègements dont le nombre et l'intensité des exemptions sont graduellement croissants avec les difficultés économiques de la zone. Parallèlement au zonage de ces quartiers prioritaires présentant des difficultés économiques et urbaines, les *zones de revitalisation rurale* (ZRR) ont été définies. Elles visent, quant à elles, à redynamiser les territoires ruraux les plus fragiles, c'est-à-dire avec une faible densité démographique et un handicap structurel sur le plan socioéconomique. Les créations d'emplois dans les ZRU et les ZRR peuvent bénéficier d'une exonération totale des cotisations patronales de sécurité sociale jusqu'à 1,5 Smic puis dégressive jusqu'à 2,4 Smic. Plus récemment, des régimes d'exonérations plus généreux ont vu le jour pour pallier les déséquilibres économiques de nouvelles zones. A ce titre, il existe depuis 2007 une franchise de cotisations sociales patronales sur la partie des salaires correspondant à 1,4 Smic horaire en faveur des emplois créés dans les *bassins d'emplois à redynamiser* (BER). Ces BER, situés aujourd'hui en Champagne-Ardenne et en région Midi-Pyrénées, se caractérisent par un taux de chômage élevé et une déperdition de population et d'emplois. Il en est de même depuis septembre 2009 pour les entreprises qui créent des activités dans les *zones de restructuration de la défense* (ZRD) situées dans des territoires affectés par le départ des installations militaires dans le cadre de la réorganisation des armées. Là-encore, l'exonération, qui vise à l'arrivée de nouveaux employeurs, est totale pour les cotisations patronales pendant 5 ans dans la limite de 1,4 Smic puis dégressive jusqu'à 2,4 Smic. Enfin, on peut citer le système dérogatoire spécifique aux 4 départements d'Outre-mer dont l'objectif affiché était de compenser les surcoûts liés à leur situation de région ultrapériphérique pour soutenir l'activité économique locale.

Ces mesures d'exonérations ciblées sur des territoires particuliers se sont largement inspirées des dispositifs nommés *enterprise zones* mis en place initialement au Royaume-Uni en 1981, puis largement expérimentés dans différents états des Etats-Unis. D'ailleurs, en 1994, les Etats-Unis ont lancé au niveau fédéral un programme nommé *empowerment zones* qui non seulement offrait des allègements fiscaux pour les entreprises s'implantant dans les zones visées, mais aussi institutionnalisait des transferts de fonds du gouvernement fédéral vers les Etats qui étaient impliqués financièrement dans les dispositifs de zones sensibles. Toutefois, il faut garder en

tête que même si la finalité de tous ces programmes est la même, à savoir promouvoir le développement économique local de ces zones et surtout améliorer le niveau de l'emploi local, les caractéristiques des dispositifs diffèrent largement entre les pays, et même d'un Etat à un autre pour les programmes américains : la générosité, la durée et le choix des taxes sur lesquelles portent les allègements varient entre les dispositifs et ils peuvent donc obtenir des résultats différents. Certains programmes - notamment ceux privilégiant les exonérations de charges sociales patronales - vont être par exemple davantage propices à la création d'emplois alors que d'autres vont davantage jouer sur l'investissement. En outre, les critères retenus pour considérer une zone géographique comme prioritaire peuvent, selon les dispositifs, privilégier soit les zones les plus dynamiques, soit au contraire les zones les plus défavorisées. Il est donc très difficile de tirer un bilan global de l'efficacité économique de ces politiques.

L'évaluation statistique des effets propres des mesures d'exonérations territoriales prises individuellement est aussi délicate. Elle doit résoudre un problème de données manquantes. Par exemple, il est possible de savoir si le nombre d'emplois a augmenté après la mise en place d'une mesure géographiquement ciblée. En revanche, il est difficile de déterminer si cette augmentation est due à la mesure en elle-même : pour cela, il faudrait connaître les augmentations d'emploi qui auraient eu lieu en l'absence de la mesure. Ainsi, pour savoir dans quelle mesure les politiques ciblées ont amélioré le développement économique local des zones visées, on aurait aimé comparer l'activité économique des territoires concernés après la mise en place de la mesure avec ce qu'aurait été leur situation en l'absence de classification en zones prioritaires. Plus précisément, pour évaluer l'influence du dispositif T sur la variable d'intérêt Y caractérisant l'activité économique (nombre d'emplois créés, nombre d'installations d'établissements, durées de chômage, *etc*), on considère que chaque zone i a virtuellement 2 valeurs potentielles de Y : Y_{i1} si la zone i est bénéficiaire du dispositif ($T = 1$, traitement) et Y_{i0} si elle n'en bénéficie pas ($T = 0$, pas de traitement). Pour chaque zone, on peut alors définir l'effet propre (ou causal) du dispositif Δ_i à partir de ces deux résultats potentiels : $\Delta_i = Y_{i1} - Y_{i0}$. Or, le problème est que l'on n'observe pas simultanément Y_{i1} et Y_{i0} : pour une zone bénéficiaire, on observe Y_{i1} mais pas Y_{i0} et inversement pour les non-bénéficiaires. L'évaluation des politiques ciblées géographiquement nécessite donc de reconstituer l'information manquante de la situation alternative inobservée. En conséquence, les méthodes d'évaluation cherchent au prix d'hypothèses plus ou moins fortes et plus ou moins valides selon les contextes à constituer un ensemble de zones non traitées aussi proches que possibles de celles qui l'ont été, appelé *groupe de contrôle*. La comparaison des résultats Y de ces deux groupes permet alors d'obtenir une approximation de l'impact de la mesure.

Toute la difficulté des évaluations statistiques des politiques publiques réside ainsi dans le fait de trouver un bon groupe de contrôle, c'est-à-dire représentatif de la situation des bénéficiaires en l'absence de dispositifs. Les véritables travaux sur les mesures d'exonérations territoriales ont commencé à émerger à partir du milieu des années 1990 en se concentrant sur une multitude d'indicateurs du marché du travail. Toutefois, ils partagent, pour la plupart, tous le même point de départ. Premièrement, ils cherchent à estimer l'effet moyen de la mesure sur ceux qui en ont effectivement bénéficié, soit *l'effet du traitement sur les traités* : $\Delta = E(Y_{i1} -$

$Y_{i0}|T_i = 1) = E(Y_{i1}|T_i = 1) - E(Y_{i0}|T_i = 1)$. Deuxièmement, ils partent du constat que la définition des territoires ciblés n'étant absolument pas aléatoire, il n'est pas pertinent d'utiliser comme groupe de contrôle l'ensemble des territoires non classés en zones prioritaires, soit en d'autres termes d'approximer $E(Y_{i0}|T_i = 1)$ qui est non observé par $E(Y_{i0}|T_i = 0)$, en raison de *biais de sélection*. En effet, les zones bénéficiant de tels dispositifs sont souvent sélectionnées sur la base d'un classement utilisant des indices synthétiques de l'activité économique et de la population. Par exemple en France, la caractérisation des zones urbaines sensibles repose sur un indice agrégé tenant compte de la population de la zone, de son taux de chômage, de la proportion de jeunes, de la proportion de travailleurs non qualifiés et du "potentiel fiscal" de la zone. En conséquence, les zones bénéficiant des politiques ciblées et celles n'en bénéficiant pas n'ont pas les mêmes caractéristiques. La comparaison de ces deux groupes ne mesure pas uniquement l'effet propre du dispositif, mais il est entaché par un "effet de population" reflétant les différences de caractéristiques entre les deux groupes. A cet "effet de population" peut s'ajouter un "biais de tendance spécifique" ou de "conjoncture" : les politiques d'exonérations territoriales viennent soutenir des zones qui, au-delà de leurs caractéristiques spécifiques fixes dans le temps, peuvent être soit en phase d'expansion soit en phase de déclin. Ainsi, si les zones traitées et non traitées se distinguent en termes d'évolution, leur simple comparaison n'est pas adéquate. Enfin, la délimitation géographique exacte de la zone ciblée est largement influencée par les pressions exercées par les élus locaux et les représentants de l'Etat. Seuls les élus locaux qui croyaient en l'efficacité des dispositifs ont joué de leur influence pour faire qualifier leurs circonscriptions en zones prioritaires. Ainsi, c'est probablement les zones pour lesquelles le dispositif est supposé *a priori* avoir les meilleures chances de retombées positives, qui ont été choisies comme territoires ciblés. Pour toutes ces raisons, les zones non-bénéficiaires ne sont pas représentatives des zones bénéficiaires du traitement : le fait d'avoir été désignés territoires ciblés est probablement lié aux résultats qu'ils pouvaient escompter avec ou sans traitement. Cette non-indépendance entre la désignation et le résultat génère que le terme $E(Y_{i0}|T_i = 1) - E(Y_{i0}|T_i = 0)$ n'est pas nul dans l'expression suivante et correspond au biais lié à ces effets de sélection :

$$E(Y_{i1}|T_i = 1) - E(Y_{i0}|T_i = 0) = \underbrace{[E(Y_{i1}|T_i = 1) - E(Y_{i0}|T_i = 1)]}_{\text{effet propre de la mesure}} + \underbrace{[E(Y_{i0}|T_i = 1) - E(Y_{i0}|T_i = 0)]}_{\text{biais de sélection}} \quad (1)$$

La littérature sur l'évaluation des politiques d'exonérations territoriales s'est quasi-exclusivement concentrée à proposer différentes méthodes pour résoudre ces problèmes de biais de sélection. Ces méthodologies se distinguent par les hypothèses sous lesquelles on peut estimer correctement l'impact causal de la mesure et l'utilisation de l'une par rapport à une autre est souvent tributaire des données disponibles. Néanmoins, l'interprétation des résultats de ces travaux nécessite que l'on s'interroge à la fois sur leur *validité interne* et leur *validité externe*, même si la littérature est restée plus silencieuse jusqu'à présent sur ces sujets. D'une part, la pertinence des hypothèses sous-jacentes à ces méthodologies mérite d'être examinée (validité interne). D'autre part, avant toute préconisation de politiques publiques, il faut se demander si les résultats obtenus sur des données particulières, une population ou une période spécifiques

sont généralisables. Par exemple, la validité externe de ces études pourraient être compromises par le fait qu'elles ne calculent en général qu'un impact causal moyen alors que l'effet d'une mesure peut différer selon les zones. L'idéal en présence d'hétérogénéité de l'effet de traitement serait de faire une évaluation de l'ensemble de la distribution et non uniquement de la moyenne, ce qui est beaucoup plus exigeant. En outre, la plupart des études font l'hypothèse implicite qu'il n'existe pas d'externalités, c'est-à-dire que la désignation des zones prioritaires n'a d'effet que sur le résultat de ces zones, et donc n'affecte pas le résultat des autres zones. Or dans notre contexte des politiques d'exonérations territoriales, on peut craindre par exemple que les créations d'entreprises dans une zone ciblée se réalisent au détriment de celles des zones voisines non traitées. Selon l'ampleur du dispositif, ces effets de bouclage peuvent intervenir et compromettre la validité externe des résultats obtenus. Il faudra donc garder en tête ces limites lors des interprétations des études d'évaluation statistique.

Même si la littérature sur les évaluations des effets propres de ces dispositifs est devenue relativement abondante pour les cas américains, elle reste encore limitée à ce jour pour le cas français. Or, au vu des différences de dispositifs, les résultats des expériences américaines sont peu extrapolables à la problématique française. En outre, ils sont mitigés : il n'y a pas de parfait consensus de leurs effets positifs sur l'accès à l'emploi. Néanmoins, les évaluations réalisées sur les EZ américaines sont riches d'enseignements du point de vue méthodologique. Le premier objectif de cet article est donc de recenser leurs méthodes économétriques pour évaluer leurs politiques ciblées sur les territoires, en en présentant leur contexte, leurs avantages et leurs limites (section 2). Le but est de voir ensuite quelles sont celles qui ont été ou qui pourraient être les plus adéquates pour le cas français. En particulier, la section 3 montre comment les évaluations de mesures similaires en France s'en sont inspirées et comment elles pourraient le faire dans de futures recherches pour les dispositifs pour lesquels aucune évaluation statistique n'existe encore.

2 Méthodes d'évaluation statistique des politiques ciblées sur les territoires : l'expérience américaine

2.1 Le principe des estimateurs de différences de différences

De manière générale, le principe d'évaluation d'une politique publique consiste à estimer un *contrefactuel*, la situation qui aurait prévalu en l'absence du dispositif que l'on souhaite évaluer. Comme évoqué précédemment, comparer les résultats des bénéficiaires d'une mesure avec ceux des non-bénéficiaires risque de donner des estimations biaisées par la présence d'effets de sélection. Comparer les variables d'intérêt à ce qu'elles étaient avant la mise en place du dispositif (estimateur avant/après) n'est pas plus satisfaisant. De nombreux facteurs tels qu'une conjoncture économique plus ou moins favorable, sont susceptibles d'expliquer des variations de résultats dans le temps, indépendamment des effets du dispositif. C'est pourquoi, les études d'évaluation de politiques publiques ont souvent utilisé la *méthode de différences de différences*

(DD), qui mélange ces deux approches. Cette terminologie vient du fait que l'on considère la variation dans le temps de la variable d'intérêt, ce qui constitue la première soustraction, et qu'on la compare entre les bénéficiaires et les non-bénéficiaires du traitement, ce qui implique une seconde soustraction. Dans notre contexte, cette méthode consisterait à comparer l'écart des performances économiques entre zones traitées et non traitées avant la mise en place du dispositif au même écart après l'application de la mesure [Cameron & Trivedi (2005)] :

$$\Delta = DD = \frac{1}{N_T} \sum_{i \in T} (Y_{it'} - Y_{it}) - \frac{1}{N_C} \sum_{j \in C} (Y_{jt'} - Y_{jt}) \quad (2)$$

où N_T est le nombre d'observations issues des zones traitées i , N_C le nombre d'observations issues des zones non-traitées j , et où t et t' caractérisent respectivement les périodes d'observation avant et après la mise en place de la politique.

L'une des différences permet ainsi d'éliminer les différences systématiques (effet de population) entre le groupe des bénéficiaires et les autres. L'autre différence permet d'éliminer l'évolution temporelle supposée identique pour les deux groupes en l'absence du dispositif. Cet estimateur de différences de différences doit théoriquement éliminer les effets de composition liés à la sélection et ne mesurer que l'impact propre, causal, de la mesure, mais au prix de l'hypothèse forte que la variable d'intérêt potentielle des deux groupes aurait effectivement évolué de manière identique en l'absence du dispositif. Par conséquent, la validité interne de cette méthode d'évaluation dépend fortement du groupe de contrôle choisi. Avec ce cadre, le problème de l'évaluation statistique consiste à déterminer un groupe de contrôle plausible, c'est-à-dire susceptible d'avoir effectivement connu des évolutions temporelles comparables à celles du groupe des bénéficiaires. Certains auteurs ont alors privilégié de caractériser finement un groupe de contrôle via des méthodes d'appariement pour optimiser les chances que cette condition soit vérifiée (section 2.2.2) alors que d'autres restent sur des groupes de contrôle *ad hoc* plus généraux mais utilisent des modélisations économétriques pour purger les différences de tendance potentielle entre les deux groupes qui ne sont pas à attribuer au dispositif (section 2.2.1). Enfin, des études plus récentes ont préféré combiner les deux approches pour caractériser leur groupe de contrôle (section 2.2.3).

En pratique, dans le cadre général de plusieurs périodes, l'estimateur de différences de différences de l'équation (2) peut aussi être obtenu par régression du modèle suivant [Givord (2010)] :

$$y_{it} = e_t + e_g + \delta T_{it} + u_{it}, \quad (3)$$

où T_{it} est une indicatrice qui vaut 1 si la zone i a été traitée à la date t ; les e_t sont des effets fixes temporels qui captent les évolutions de la conjoncture affectant les deux groupes et e_g est l'effet fixe de groupe rendant compte des différences permanentes entre les groupes. δ correspond alors à l'impact causal du dispositif. Pour rendre ce modèle opérationnel, il suffit par exemple de mettre une indicatrice pour chacune des périodes (cela capte e_t), une indicatrice de groupe

($e_g = 1_{T_i=1}$) et une indicatrice croisée pour indiquer si la zone i a bénéficié d'un dispositif ciblé à la date t . Lorsque l'on dispose de données de panel, il est pertinent d'utiliser l'information supplémentaire qu'elles apportent pour mieux contrôler l'hétérogénéité individuelle entre zones avant la mise en place du traitement [Cameron & Trivedi (2005)]. Cela revient à remplacer l'effet moyen de groupe par des effets fixes individuels :

$$y_{it} = e_t + \alpha_i + \delta T_{it} + u_{it}. \quad (4)$$

On peut alors estimer δ par des techniques standard d'économétrie des données de panel, comme des estimations en différences (certains parlent alors d'estimateurs des *triples différences* ou *difference in difference in difference* (DDD)) ou des estimateurs à effets fixes (*within*). Ces deux procédures permettent d'éliminer les effets fixes α_i , soit en différenciant les variables selon la dimension temporelle ou en utilisant les écarts à la moyenne temporelle.

Dans le contexte des politiques d'exonérations territoriales, deux formes particulières du modèle à effets fixes de l'équation (4) ont été estimées. La première consiste à modéliser l'effet temporel par une tendance linéaire :

$$y_{it} = \beta.t + \alpha_i + \delta T_{it} + u_{it}. \quad (5)$$

A aussi été proposée une version plus raffinée qui autorise la variable dépendante à suivre une tendance temporelle spécifique dans chaque zone. Ce modèle particulier à effets fixes s'appelle le *modèle à croissance aléatoire* [Heckman & Hotz (1989)] et s'écrit comme suit :

$$y_{it} = \sigma_t + \beta_i.t + \alpha_i + \delta T_{it} + u_{it} \quad (6)$$

où σ_t est un effet temporel fixe qui est modélisé par une indicatrice pour chaque année et où β_i mesure la tendance spécifique à chaque zone i .

2.2 Caractérisation du groupe de contrôle et résultats

2.2.1 Choix d'un groupe de contrôle *ad hoc*

Les premières évaluations des politiques d'exonérations territoriales qui ont appliqué une méthode de différences de différences ont défini leur groupe de contrôle de manière *ad hoc*. Elles ont comparé la situation des bénéficiaires avec un groupe considéré comme semblable, mais il existe toujours une part d'arbitraire dans ce choix.

Papke (1994) a été la première à proposer une évaluation économétrique des politiques d'exonérations territoriales et plus précisément de l'impact sur l'emploi et l'investissement des *enterprise zones* (EZ) mises en place dans l'Indiana en 1983. Dans ce dispositif, les zones étaient sélectionnées sur la base de leur taux de chômage et de leur taux de pauvreté et elles donnaient droit à des exonérations de taxes locales, à des crédits d'impôts à l'embauche (limités

à 10% du salaire) et les salariés bénéficiaient d'un crédit d'impôt sur le revenu. Papke (1994) exploite la dimension panel de ses données et estime les modèles des équations (5) et (6) via l'estimateur *within* et à partir de régressions en différences. Elle effectue son étude au niveau des juridictions fiscales, dont la taille est approximativement celle d'une EZ. Par conséquent, pour la construction de sa variable de traitement T_{it} , elle utilise comme groupe de contrôle toute juridiction fiscale de l'Indiana ne contenant pas une EZ, soit l'ensemble des juridictions non-bénéficiaires. Toutefois, les modélisations économétriques utilisées permettent une comparaison des deux groupes toutes choses étant égales par ailleurs, c'est-à-dire à caractéristiques fixes dans le temps et à effets temporels similaires. Papke (1994) conclut que les allègements circonscrits à une zone stimulent l'emploi et l'investissement local. En particulier, la classification en EZ en Indiana aurait fait diminuer le chômage local annuel de 19%.

Boarnet & Bogart (1996) pour le programme d'EZ du New Jersey, suivis de Moore (2003) et de Neumark & Kolko (2010) pour le dispositif d'EZ mis en place en Californie, se sont fortement inspirés de la méthodologie de Papke (1994). Toutefois, leurs travaux utilisent des groupes de contrôle plus restreints que les seules zones non-bénéficiaires afin d'essayer de s'affranchir du problème d'endogénéité selon lequel il existerait des facteurs d'hétérogénéité inobservée entre zones qui joueraient simultanément sur la décision de classification en territoires ciblés et sur les outputs potentiels. Pour que le paramètre δ capte au mieux l'impact de l'EZ, Boarnet & Bogart (1996) utilisent comme groupe de contrôle les seules municipalités qui ont candidaté pour le dispositif d'EZ mais qui finalement n'ont pas été retenues. Ils considèrent que ces zones qui étaient candidates sont un meilleur groupe de contrôle que l'ensemble des zones non-EZ non seulement car elles ont des caractéristiques socioéconomiques plus proches, notamment en termes de taux de chômage, pauvreté, composition de la population - elles remplissaient aussi les critères d'éligibilité - et car comme les zones désignées EZ, elles en attendaient *a priori* des retombées positives. Moore (2003) et Neumark & Kolko (2010) évoquent ces mêmes arguments pour limiter leur échantillon d'estimation aux zones qui ont été initialement EZ en 1986 (groupes des traités) et à celles qui l'ont été lors d'une seconde vague du programme en 1991-1992 (groupe de contrôle).

Alors que le dispositif d'EZ du New Jersey offrait pour toute entreprise s'implantant ou s'agrandissant dans une zone ciblée des crédits d'impôts à l'embauche, des allègements dans les charges d'assurance chômage ou encore des réductions de taxes sur les ventes au détail et certains achats tels que les matériaux de construction, Boarnet & Bogart (1996) ne trouvent aucun impact significatif sur le nombre de créations d'emplois dans les municipalités ayant des EZ. Moore (2003) trouve des résultats opposés pour la Californie. Le nombre d'entreprises dans les ZIP (code ZIP : *Zooning Improvement Plan*) contenant une EZ désignée en 1986 aurait augmenté d'environ 5% de plus que dans le groupe de contrôle. Des analyses plus détaillées montrent que cet impact moyen cache en fait de fortes disparités en fonction de la taille et du secteur des entreprises et plaident en faveur d'une hétérogénéité importante de l'impact du dispositif. Par exemple, les hausses du nombre d'entreprises auraient été plus marquées dans le secteur de la finance, de l'assurance et de l'immobilier et dans celui du commerce de détails.

Toutefois, ces résultats sont à prendre avec précaution car ces deux études utilisent des données agrégées, respectivement au niveau des municipalités pour Boarnet & Bogart (1996) et au niveau des ZIP pour Moore (2003). Pourtant les contours géographiques des EZ ne correspondent qu'approximativement à ces niveaux d'agrégation : les EZ ne couvrent souvent que partiellement le territoire d'une municipalité ou d'une ZIP. Ces approximations sur l'identification des zones géographiques introduisent des erreurs de mesure : on désigne des zones appartenant à une EZ alors qu'elles ne le sont pas et inversement. Or, si les unités géographiques sont incorrectement classées comme EZ ou groupe de contrôle, il y a de forts risques d'avoir un biais dans l'évaluation de l'impact de l'EZ sur le développement économique local. D'ailleurs, Boarnet & Bogart (1996) avancent que l'absence d'effet significatif dans leur étude pourrait provenir d'un transfert d'emplois s'opérant entre les EZ et les non-EZ mais au sein même des municipalités concernées.

Pour pallier ce problème d'identification des zones ciblées, Neumark & Kolko (2010) utilisent des méthodes cartographiques pour définir précisément (rue par rue) les contours des EZ en Californie et donc créer une variable de traitement T_i moins sujette aux erreurs de mesure. Ils disposent aussi de données sur la période 1992-2004 leur donnant les adresses exactes des établissements pour chaque année et en tirent une information localisée très fiable sur le niveau de l'emploi et le nombre d'établissements. Comme Moore (2003), ils privilégient pour définir le groupe de contrôle des EZ initiales en Californie les zones qui le sont devenues par la suite. Les résultats ainsi obtenus contredisent ceux de Moore (2003) puisque Neumark & Kolko (2010) trouvent que l'impact des EZ sur l'emploi est non significatif à la fois à court terme et à long terme.

Même si les résultats de Neumark & Kolko (2010) paraissent plus robustes que ceux de Moore (2003) en raison de la définition plus précise des EZ, une même critique peut être avancée à ces deux études. Choisir comme groupe de contrôle les futurs zonages ne va pas sans poser des problèmes [Ham, Swenson, Imrohoroglu & Song (2011)]. C'est le cas lorsqu'est adoptée une chronologie de désignation liée aux difficultés des zones. Par exemple, si les autorités ont désignés pour la première vague des EZ celles qui avaient le plus besoin d'aide, les futures zones bénéficiaires du dispositif sont moins en difficultés que les premières et donc l'indépendance des critères de sélection des zones et des outputs potentiels, que l'on cherchait à obtenir par le choix d'un tel groupe de contrôle, est aussi violée. On obtiendrait aussi une estimation biaisée si la désignation comme EZ initiale était basée sur le fait que la zone ait connu récemment un choc négatif. C'est pourquoi, Ham et al. (2011) privilégient d'autres groupes de contrôle *ad hoc* et testent via une approche par variables instrumentales si l'indépendance entre les variables d'intérêt et de traitement est vérifiée pour ces groupes. L'originalité de leur étude est en fait triple. Premièrement, ils étudient l'impact des EZ non pas sur un Etat américain particulier, mais au niveau national en considérant l'ensemble des EZ de tout le pays. Ils mobilisent des données à l'échelle de l'unité de recensement (le *census tract*, comptant en moyenne 5 000 habitants) et construisent leur variable de traitement comme suit : une unité de recensement (*census tract*) est considérée comme bénéficiaire de l'EZ si 50% ou plus de cette zone est couverte par le dispositif d'EZ. Deuxièmement, ils caractérisent trois groupes de

contrôle *ad hoc* et déterminent le meilleur non pas de façon arbitraire mais au moyen de tests effectués sur les données. En particulier, ils considèrent successivement trois manières de définir leur groupe de contrôle : i) ils ne gardent dans le groupe de contrôle qu’un census tract par unité bénéficiaire, l’unité non traitée la plus proche ; ii) ils gardent dans le groupe de contrôle pour chaque unité bénéficiaire toutes les zones non-traitées du même état, mais qui lui sont contigües ; iii) ils élargissent leur groupe de contrôle à toutes les zones non bénéficiaires au sein de l’état. Troisièmement, Ham et al. (2011) innovent en regardant si l’hypothèse identifiante des estimateurs de différences de différences selon laquelle quand on contrôle pour les effets fixes et les tendances temporelles les groupes des traités et de contrôle choisi sont parfaitement comparables, est vérifiée. Pour contrôler l’endogénéité potentielle qui pourrait subsister, ils instrumentent la variable de traitement T_i en utilisant les valeurs retardées de l’output. Ils vérifient alors que leurs résultats sont peu modifiés selon qu’ils mettent ou non en œuvre la méthode par variable instrumentale. Cela permet de lever les doutes sur la robustesse de l’hypothèse identifiante de leur stratégie d’estimation de différences de différences.

Ham et al. (2011) examinent l’impact des EZ sur plusieurs indicateurs du marché du travail local. Ils trouvent que le dispositif des EZ a généré une réduction importante du taux de chômage local (-1,6%) et du taux de pauvreté (-6,1%), a accru l’emploi et augmenté les salaires et les revenus des habitants des zones bénéficiaires. Avec une méthodologie similaire, ils évaluent les retombées du programme fédéral américain d’*Empowerment Zones* et arrivent à la conclusion que cette mesure a aussi eu des répercussions positives sur l’ensemble des indicateurs : baisses significatives du taux de chômage et de pauvreté, augmentation de l’emploi et des salaires des résidents de ces territoires ciblés, mais dans une moindre mesure que les programmes d’EZ.

Comme il n’est pas toujours évident de trouver de bonnes variables instrumentales pour la variable de traitement, l’hypothèse identifiante est souvent non testée et il plane alors le doute d’une endogénéité potentielle. C’est pourquoi, d’autres auteurs ont renoncé au choix d’un groupe de contrôle *ad hoc* pour privilégier la méthode par appariement.

2.2.2 Détermination d’un groupe de contrôle par la méthode d’appariement

L’idée des méthodes d’appariement est de définir de façon plus précise, à partir de caractéristiques observables pré-définies, un groupe de contrôle au lieu de raisonner avec les zones non traitées dans leur ensemble, celles pas encore désignées mais qui le seront ultérieurement, ou encore les zones géographiquement proches. Pour réduire le problème d’endogénéité, lié à l’existence d’une sélection probable dans les dispositifs, les méthodes d’appariement vont contrôler en partie l’effet de sélection en contrôlant des différences observables entre les bénéficiaires et les non-bénéficiaires. Elles consistent à apparier chaque zone traitée à une zone “jumelle” non traitée possédant les mêmes caractéristiques. Ces méthodes reposent sur l’hypothèse centrale très forte d’*indépendance conditionnelle* qui postule que les différences observables entre les bénéficiaires et les non-bénéficiaires captent l’ensemble des déterminants de la sélection des

bénéficiaires. En d’autres termes, en dehors des observables choisies, il n’existe pas d’autres caractéristiques qui influent à la fois les outputs potentiels et le choix du traitement. Plus précisément dans notre contexte des politiques d’exonérations territoriales, cette hypothèse signifie que le fait pour deux zones présentant des caractéristiques proches d’être pour l’une bénéficiaire d’un tel dispositif et pour l’autre pas, n’est pas lié à l’anticipation des effets escomptés de la zone. Le risque que cette hypothèse d’indépendance conditionnelle ne soit pas vérifiée est important ; mais elle ne peut malheureusement pas être facilement testée.

Les méthodes d’appariement sont aussi sujettes à la condition dite d’*existence d’un support commun* : pour pouvoir comparer une zone traitée avec une zone non-traitée identique, encore faut-il qu’elle existe. Si pour certaines caractéristiques X les zones étaient certaines de bénéficier du dispositif d’exonérations, il ne serait pas possible de trouver une zone non-traitée comparable. Dans l’exemple des EZ, c’est le fait que la création des zones ne soit pas déterminée exactement par les caractéristiques socio-démographiques qui permet de les utiliser comme variables de conditionnement. Si cela n’était pas le cas, il ne serait pas possible de trouver des territoires comparables aux EZ. Ainsi, les méthodes d’appariement ne sont valables que si pour chaque zone traitée, on peut trouver une observation du groupe de contrôle assez proche. Les zones pour lesquelles cette propriété n’est pas vérifiée doivent alors être exclues de l’analyse.

En pratique, plusieurs méthodes d’appariement ont été utilisées pour évaluer des politiques publiques [Givord (2010)]. La méthode la plus simple consisterait à comparer chaque zone avec sa “jumelle”. Toutefois, il peut être rare de trouver une zone du groupe de contrôle ayant exactement les mêmes caractéristiques observables, donc on choisit le “plus proche voisin”. Cette technique d’appariement par plus proche voisin peut être faite sans remise (une zone du groupe de contrôle ne peut être appariée qu’une fois avec une zone du groupe de traitement) ou avec remise (une zone non-traitée peut servir de “jumelle” pour plusieurs zones traitées). Elle a néanmoins l’inconvénient que la qualité de l’appariement n’est pas contrôlée : certains plus proches voisins peuvent se révéler être très différents des zones traitées. Or, cette méthode traite de la même manière les paires proches et moins proches. Par ailleurs, appairer à une seule zone prive de l’information apportée pour toutes les autres, ce qui réduit *a priori* la précision de l’estimation : il est dommageable si une zone bénéficiaire a plusieurs zones “jumelles” très proches de n’en choisir qu’une seule.

C’est pourquoi, les travaux d’évaluation des politiques d’exonérations territoriales ont utilisé des variantes de cette technique d’appariement. Par exemple, plutôt que de se limiter au plus proche voisin, Bostic & Prohofsky (2006) choisissent d’appairer avec un nombre fixe de 3 plus proches voisins. Plus précisément, Bostic & Prohofsky (2006), à partir de données de l’administration fiscale, cherchent à voir si les revenus des individus participant à une EZ en Californie sont significativement plus élevés que ceux du groupe de contrôle, puisque les EZ en Californie concèdent des exonérations de taxes et des crédits d’impôts sur les salaires (à la hauteur de 50% du salaire des cinq premières années). Leur appariement est effectué selon les trois critères suivants : le statut matrimonial, le nombre de personnes à charge et l’appartenance à une ZIP comparable. Pour chaque individu bénéficiant du dispositif de l’EZ (notamment en termes de réduction d’impôts), Bostic & Prohofsky (2006) cherchent les individus parmi les

non-traités qui ont exactement le même statut matrimonial, le même nombre de personnes à charge, et qui appartiennent aux 5 ZIP les plus proches de la sienne en termes de composition ethnique, pauvreté et revenus. Ces individus sont classés selon leur degré de similarité en termes des salaires, âge et crédits d’impôts et seuls les 3 mieux classés sont définitivement retenus pour constituer le groupe de contrôle. A partir d’un tel appariement, Bostic & Prohofsky (2006) montrent que les revenus des individus bénéficiant du dispositif des EZ augmenteraient plus rapidement que ceux des individus du groupe de contrôle.

Pour répondre aux limites de la méthode d’appariement du plus proche voisin, une autre variante va plus loin en utilisant pour l’appariement de chaque zone traitée toutes les zones du groupe de contrôle, mais en les pondérant en fonction de leur distance à la zone traitée : le contrefactuel est ainsi calculé par une estimation à *noyau* (*kernel*). On fixe alors la taille de la fenêtre (bandwidth) du noyau. Comme en dehors de cette fenêtre les poids sont très faibles, plus la fenêtre est petite, plus le contrefactuel d’une zone bénéficiaire ne prend en compte que les zones du groupe de contrôle dont les caractéristiques observables sont très proches de celle-ci.

Toutefois, pour que l’hypothèse d’indépendance conditionnelle soit vérifiée, il est souhaitable d’utiliser le maximum d’informations et donc d’apparier selon un grand nombre de variables de conditionnement. Mais plus ce nombre de variables est élevé, plus il est difficile de trouver un voisin proche et donc plus les estimateurs risquent d’être biaisés. C’est pourquoi, la plupart des évaluations empiriques, et notamment celles dans notre contexte, ont préféré *l’appariement sur le score de propension*.

La méthode d’appariement par score de propension (*propensity score matching*) consiste à modéliser la probabilité d’être traité en fonction de différentes variables de contrôle et de considérer que les zones qui ont la même probabilité d’être traitées sont comparables et constituent un bon groupe de contrôle. L’appariement se fait ainsi à partir d’un indicateur unique, cette probabilité, appelée aussi *score de propension*. Cette probabilité qu’une unité géographique de caractéristiques X soit traitée est en général estimée par une régression logit ou probit. Chaque unité traitée est ensuite appariée avec les unités non traitées ayant les scores les plus proches possibles ; l’ensemble de ces unités non traitées constituant le groupe de contrôle. Pour que l’hypothèse de support commun soit vérifiée, il ne faut pas que les variables de conditionnement choisies expliquent “trop bien” le fait d’être traité, car cela peut rendre impossible un appariement avec les zones bénéficiaires très proches. Il est nécessaire que les zones qui font face au même score de propension soient présentes dans les deux sous-populations (bénéficiaires et non bénéficiaires de l’exonération). De nouveau, des algorithmes d’appariement sur ce score de propension plus ou moins raffinés ont été proposés dans la littérature. Par exemple, dans O’Keefe (2004), chaque EZ est appariée à la zone non-EZ avec le score le plus proche parmi les zones non ciblées du même canton. Une même zone non-EZ peut donc constituer le meilleur appariement pour plusieurs EZ ; il s’agit d’un appariement avec remise. Les variantes de cette approche assignent un poids aux unités non traitées qui varie en fonction de la différence de score de propension entre zones traitées et non traitées [Caliendo & Kopienig

(2008)]. L'appariement par noyau peut être par exemple appliqué en utilisant l'estimation du score de propension comme mesure de la distance entre deux observations.

La mise en œuvre de l'estimateur de différences de différences avec appariement [Heckman & Smith (1997)] nécessite une légère modification de l'équation (2) comme suit :

$$\Delta = WDD = \frac{1}{N_T} \sum_{i \in T} (Y_{it'} - Y_{it}) - \frac{1}{N_C} \sum_{j \in C} w_{ij} (Y_{jt'} - Y_{jt}) \quad (7)$$

où w_{ij} est un poids appliqué aux observations des zones non traitées, qui tient compte des scores de propension et de leur différence par rapport à ceux des zones traitées. Il s'agit d'un estimateur de *différences de différences pondéré* (Weighted difference-in-difference, WDD). De nouveau, les poids peuvent être déterminés selon différents algorithmes d'appariement. On peut là-encore obtenir cet estimateur via des régressions semblables aux équations (3) à (6) en restreignant l'échantillon d'estimation aux zones traitées et à leurs "jumelles" et en définissant la variable de traitement T_i à partir de ces deux groupes.

Avec cette technique d'appariement par score de propension, O'Keefe (2004) étudie l'effet sur l'emploi de 39 EZ en Californie en s'intéressant à deux problématiques nouvelles par rapport à Moore (2003) : elle regarde d'une part si l'effet des dispositifs sur l'emploi serait plutôt transitoire ou permanent, thématique reprise ultérieurement par Neumark & Kolko (2010) et d'autre part si les crédits d'impôts proposés sur les salaires sont captés par les salaires les plus élevés, thématique reprise elle ultérieurement par Bostic & Prohofsky (2006). En utilisant des données annuelles d'emploi au niveau des établissements et après un appariement réalisé à l'échelle de l'unité de recensement (le *census tract*), O'Keefe (2004) trouve au final un effet significatif des EZ de 3% sur la croissance de l'emploi au cours des six premières années¹. Mais cet effet est seulement transitoire. Elle suggère que la disparition de l'effet à moyen terme pourrait être liée au processus d'incitations financières à l'embauche limitées à cinq ans et aux possibilités d'emplacements disponibles dans la zone pour l'implantation de nouvelles entreprises qui se réduisent au fil des années. Pour les salaires, O'Keefe (2004) ne trouve pas d'impact significatif du dispositif d'EZ.

Toutefois, ces résultats ont été remis en cause par les travaux de Neumark & Kolko (2010) et Bostic & Prohofsky (2006) et fortement critiqués. Premièrement, Neumark & Kolko (2010) reprochent à O'Keefe (2004) l'utilisation de données agrégées au niveau du *census tract*. Les contours géographiques des EZ en Californie ne correspondent qu'approximativement à ceux des *census tracts*. Il y a donc un problème d'identification des zones ciblées qui peut générer un biais de l'estimation. Deuxièmement, Neumark & Kolko (2010) critiquent le fait que l'appariement par score de propension de O'Keefe (2004) soit réalisé à partir d'observations post-traitement. Or, un tel appariement peut masquer une partie des effets de l'EZ puisqu'il conditionne sur des variables qui ont pu également être affectées par le dispositif. Cela peut créer un

1. Il est à noter que contrairement aux modèles de régression de différences de différences, O'Keefe (2004) n'introduit pas d'effets temporels fixes dans son modèle. Son modèle de régression correspond plus à une simple différence qu'à une double différence.

problème d'endogénéité. Il est ainsi davantage recommandé de procéder à l'appariement sur des observations pré-traitement mais faut-il encore qu'elles soient disponibles. Troisièmement, ils soulignent le biais potentiel lié à une éventuelle corrélation entre des caractéristiques inobservables jouant simultanément sur la croissance de l'emploi et la probabilité d'appartenir à une EZ. En effet, l'*hypothèse d'indépendance conditionnelle*, qui postule qu'il n'existe pas de variables inobservables qui interviennent à la fois dans le processus de sélection des zones traitées et dans les performances économiques de celles-ci, ne peut pas être facilement testée et n'est donc pas garantie. Enfin, on aurait pu rajouter que O'Keefe (2004) limite son évaluation statistique aux EZ pour lesquelles un bon appariement a pu être trouvé selon le critère suivant : la différence entre les scores de propension de l'EZ et sa zone non-EZ appariée devait être inférieure à 10%. Or, comme dans le cadre des dispositifs d'EZ, les territoires ciblés sont désignés sur des caractéristiques prédéfinies, cette élimination de certaines zones peut entraîner un biais dans l'estimation [Heckman, Ichimura, Smith & Todd (1998), Lorenceau (2009)].

Du fait de ces limites soulevées par l'étude de O'Keefe (2004), Neumark & Kolko (2010) ont renoncé à employer l'approche par appariement pour déterminer leur groupe de contrôle. Ils préfèrent utiliser le groupe de contrôle *ad hoc* des futurs territoires ciblés, ce qui n'est pas sans poser aussi des problèmes. En tout cas, l'ensemble de ces résultats viennent un peu plus mitiger ceux déjà obtenus par Papke (1994) et Boarnet & Bogart (1996) pour les premiers dispositifs d'EZ.

Oakley & Tsao (2006) et Busso & Kline (2008) examinent, quant à eux, l'efficacité de la première vague du programme fédéral américain d'*Empowerment Zones*. Alors que ces études utilisent toutes les deux la technique d'appariement par score de propension et les données au niveau du census tract, ils obtiennent des conclusions opposées. Plus précisément, Oakley & Tsao (2006) reprennent la méthodologie de O'Keefe (2004) en appariant chaque zone traitée à la zone ayant le score de propension le plus proche parmi les zones non ciblées dans la même ville. Toutefois, contrairement à O'Keefe (2004), ils déterminent leurs scores de propension qu'à partir de variables de conditionnement pré-traitement : ils utilisent les variables de recensement relatives aux années 1980 et 1990 alors que le dispositif d'*Empowerment Zones* (EPZ) étudié a été instauré en 1994. Avec un tel groupe de contrôle, ils ne trouvent aucun impact significatif des EPZ sur la progression des revenus et sur la réduction du chômage et de la pauvreté dans les territoires ciblés. Busso & Kline (2008) combinent l'approche de caractérisation *ad hoc* et celle d'appariement par score de propension pour déterminer leur groupe de contrôle. Ils espèrent ainsi en mutualiser les avantages. Ils commencent par restreindre leur groupe de contrôle aux zones qui ont candidaté mais n'ont pas été choisies et à celles qui ont fini par bénéficier de ce programme lors d'une vague ultérieure, ils calculent les scores de propension, effectuent leur appariement puis appliquent l'équation (7). Ils trouvent alors que ce dispositif, sur la période 1994-2000, a eu un effet positif sur l'emploi local d'environ 4 points de pourcentage et un effet négatif sur le taux local de pauvreté d'ampleur similaire. Comme nous l'avons déjà évoqué, ces retombées positives sur l'activité locale des EPZ ont été corroborées ultérieurement par l'étude de Ham et al. (2011) qui tenait compte explicitement de la potentielle endogénéité de la variable de traitement.

2.2.3 Les régressions sur scores de propension

D'autres évaluations statistiques ont préféré utiliser les scores de propension comme variables explicatives additionnelles dans les régressions et parfois s'en servent en plus pour définir un bon groupe de contrôle.

Les résultats mitigés de Papke (1994), Boarnet & Bogart (1996), Moore (2003), O'Keefe (2004) et Neumark & Kolko (2010) laissent penser que certains programmes d'*enterprise zones* peuvent connaître une meilleure réussite que d'autres. Les conclusions obtenues à partir de l'évaluation d'un seul programme d'un Etat particulier sont difficilement généralisables à d'autres états et d'autres périodes de temps. Bondonio & Engberg (2000) et Bondonio & Greenbaum (2007) parlent alors d'un *problème de validité externe*. Ils tentent de le résoudre en plus des problèmes de biais de sélection. Ils apportent un éclairage plus général sur l'efficacité des dispositifs sur territoires ciblés en considérant respectivement cinq et dix Etats américains ayant connu des expériences d'EZ et en tenant compte de l'hétérogénéité des programmes entre les différents Etats. Ils contrôlent notamment pour la valeur monétaire des incitations de chaque programme.

Par ailleurs, Bondonio & Engberg (2000) et Bondonio & Greenbaum (2007) privilégient une méthode d'évaluation proche des modèles à croissance aléatoire mais qui intègre les scores de propension : *les régressions sur des scores de propension*. La différence principale entre ces deux familles de modèles est la façon dont les taux de croissance spécifiques aux zones sont représentés. Plus précisément, le coefficient de tendance spécifique à chaque zone i , β_i , du modèle à croissance aléatoire [équation (6)] est remplacé par une fonction linéaire du score de propension selon l'expression suivante :

$$\beta_i = \beta_j + \gamma_j \hat{P}R_i^j + v_i \quad (8)$$

avec $\hat{P}R_i^j$ la probabilité estimée que la zone i dans l'Etat j soit désignée comme EZ (même définition du score de propension que dans les méthodes précédentes) et v_i un terme d'erreur qui inclut les caractéristiques inobservables des unités géographiques.

Une autre différence entre ces régressions sur les scores de propension et le modèle de croissance aléatoire est que la variable de traitement T_i joue sur le taux de croissance de l'emploi et non sur l'emploi en niveau comme dans l'équation (6). En conséquence, les régressions sur les scores de propension consistent à estimer en première étape un modèle probit par état j expliquant le fait d'avoir été désigné ou non comme EZ en fonction de caractéristiques socioéconomiques X de la zone, et notamment celles utilisées comme critères d'éligibilité définis par la législation (taux de chômage, de pauvreté, *etc*). A l'issue des estimations, on calcule pour chaque zone i leurs probabilités prédites $\hat{P}R_i^j = \hat{P}(T_i = 1|X_i)$. En deuxième étape, le modèle suivant est estimé :

$$\Delta \ln(Y_{it}) = \sigma_t + \beta_j + \gamma_j \hat{P}R_i^j + \delta T_{it} + \varepsilon_{it}. \quad (9)$$

Bondonio & Engberg (2000) affirment que si v_i est non corrélé avec le processus de désignation des EZ, alors les régressions sur scores de propension donnent des estimateurs convergents de l'impact des dispositifs de territoires ciblés et qu'elles seront plus efficaces que les modèles à croissance aléatoire. Dans le cas contraire, les estimateurs seront non convergents. Toutefois, ils avancent plusieurs arguments crédibles en faveur de l'absence de cette corrélation.

Enfin, comme les scores de propensions ont été calculés, ils vont être utilisés pour restreindre l'échantillon afin d'avoir une meilleure évaluation de l'effet du traitement [δ dans l'équation (9)]. Dans l'approche d'appariement par score de propension, il était crucial de limiter l'échantillon aux zones avec des scores de propension pour lesquelles il y a à la fois des observations parmi les traités et le groupe de contrôle. Sans effectuer un appariement complet comme dans l'approche de l'appariement, Bondonio & Engberg (2000) et Bondonio & Greenbaum (2007) excluent de l'analyse les zones avec des scores de propension inférieurs au premier percentile de la distribution des scores des EZ et les zones avec un score supérieur au 99ème percentile de la distribution des scores des non-EZ. Cette limitation de l'échantillon sur la base des valeurs des scores de propension donne de meilleurs estimateurs de l'impact que d'utiliser l'ensemble des non-EZ comme groupe de contrôle.

Bondonio & Engberg (2000) et Bondonio & Greenbaum (2007) utilisent des données d'établissement et les agrègent au niveau des ZIP. Bondonio & Engberg (2000) constatent qu'il n'y a aucun impact significatif des EZ sur la croissance de l'emploi ou le nombre d'entreprises. Comme Boarnet & Bogart (1996), ils suspectent que ce faible effet au niveau agrégé soit à attribuer aux transferts d'emplois des zones non ciblées vers les *enterprise zones* et de vieilles entreprises vers de nouvelles entreprises et cela au sein d'un même ZIP. Par ailleurs, en ajoutant des variables d'interactions relatives aux caractéristiques des programmes (générosité, durée, choix des taxes, *etc*), ils montrent que le niveau des incitations financières pour les entreprises d'une EZ ne semble pas contribuer notablement à l'impact de l'emploi local du programme EZ.

Bondonio & Greenbaum (2007) étendent l'étude de Bondonio & Engberg (2000) en évaluant séparément l'impact des EZ sur trois catégories d'entreprises : les nouvelles entreprises dans l'EZ, les entreprises existantes préalablement dans la zone et les établissements en voie de disparition. Conformément à l'intuition de Bondonio & Engberg (2000), ils trouvent que si l'on se restreint aux entreprises nouvellement créées, l'effet du dispositif sur le taux de croissance de l'emploi est égal à 25%. Toutefois, cet impact est compensé par des pertes d'emplois accélérés dans les établissements en faillite.

2.3 Hétérogénéité du traitement et effets de bouclage

Au delà des principaux avantages et inconvénients de chaque méthode d'évaluation, il ressort que les résultats des politiques ciblées sur les territoires aux Etats-Unis (*Enterprise and Empowerment Zones*) sont mitigés. Sur données américaines, beaucoup d'études trouvent aucun effet significatif sur la croissance de l'emploi dans les zones ciblées [Boarnet & Bogart

(1996), Bondonio & Engberg (2000), Billings (2009), Hanson (2009), Neumark & Kolko (2010), Lynch & Zax (2011)], sur la probabilité de retour à l'emploi pour les chômeurs résidant dans ces territoires [Elvery (2009)], sur les salaires et sur les taux locaux de chômage et de pauvreté [O'Keefe (2004), Oakley & Tsao (2006)]. Mais d'autres trouvent des effets bénéfiques au moins à court terme non seulement sur l'emploi ou les créations d'entreprises [Papke (1994), Moore (2003), O'Keefe (2004), Busso & Kline (2008), Ham et al. (2011)], mais aussi sur les salaires [Bostic & Prohofsky (2006), Ham et al. (2011)] et sur la réduction du chômage et de la pauvreté locale [Papke (1994), Ham et al. (2011)]. Ces résultats *a priori* contradictoires laissent penser qu'il existe une forte hétérogénéité de l'impact suivant de multiples dimensions géographiques, sectorielles ou individuelles. Or la plupart de ces études se sont concentrées à évaluer un effet moyen à partir de données agrégées. Pourtant, Moore (2003) a montré des différences sectorielles dans les créations d'entreprises générées par le dispositif; Ham et al. (2011) ont conclu à l'existence de disparités d'efficacité des EZ entre Etats américains et enfin Bondonio & Greenbaum (2007) ont montré que l'impact moyen des EZ sur l'emploi masque de fortes différences selon le type d'entreprises considérées. Les politiques de territoires ciblés seraient en effet favorables à l'emploi dans les entreprises nouvellement créées mais elles accéléreraient les pertes d'emploi dans les établissements en situation difficile. L'étude de ces effets hétérogènes est une piste de recherche qui mériterait d'être approfondie afin de mieux savoir pourquoi certaines politiques ont de meilleures retombées dans certains contextes, secteurs ou certains territoires [Briant, Lafourcade & Schmutz (2012)].

Outre cette probable forte hétérogénéité de l'impact des politiques d'exonérations territoriales, Hanson & Rohlin (2012) mettent en évidence que la majorité des études d'évaluation ne tiennent pas compte des *effets de bouclage* (*spillover effects*). Même si certains auteurs évoquent l'existence d'externalités de tels dispositifs dans leurs interprétations, ils les négligent dans leur évaluation de l'impact moyen. Or, si les politiques de discrimination territoriale ont eu des répercussions sur l'activité économique de certaines zones non-ciblées, l'évaluation des programmes par le seul *effet du traitement sur les traités* n'est pas suffisante. Par ailleurs, si des zones non-traitées touchées par ces externalités appartiennent au groupe de contrôle, les effets estimés de la politique territoriale seront biaisés.

Hanson & Rohlin (2012) tentent de mesurer l'ampleur de ces externalités non seulement sur les zones non-traitées géographiquement proches des EPZ mais aussi sur celles économiquement proches. Leur idée sous-jacente est de voir si les répercussions positives du dispositif fédéral d'EPZ sur les créations d'emplois ou d'entreprises dans les territoires ciblés [Busso & Kline (2008), Ham et al. (2011)] ne se réaliseraient pas au détriment du développement économique de zones voisines non bénéficiaires des programmes d'aides. Le groupe d'intérêt dans leur évaluation n'est donc pas les zones bénéficiaires, mais les zones géographiquement ou économiquement proches des EPZ. Hanson & Rohlin (2012) vont comparer les résultats de ces dernières tout d'abord avec un groupe de contrôle *ad hoc* constitué des zones géographiquement ou économiquement proches des zones qui s'étaient portées candidates pour l'EPZ mais qui n'avaient pas été retenues. Ils réitèrent ensuite leur évaluation de différences de différences en utilisant une méthode d'appariement par score de propension pour déterminer l'impact sur

les zones économiquement similaires. Ils trouvent alors, quelle que soit la méthode, que les zones adjacentes aux EPZ ou économiquement similaires ont connu une baisse de leur nombre d'établissements et de salariés à la suite de la mise en place du dispositif. Ils démontrent même que l'utilisation de ces zones sujettes aux externalités comme groupe de contrôle pour évaluer l'effet du traitement sur les traités aurait tendance à le sur-estimer. En effet, dans beaucoup de secteurs, les externalités négatives sur les zones non-bénéficiaires proches plus que contrebalanceraient les effets positifs sur les zones bénéficiaires. L'étude récente de Hanson & Rohlin (2012) jette ainsi le doute sur la pertinence et la validité externe des travaux antérieurs centrés sur le seul effet de traitement sur les traités. La prise en compte des effets de bouclage mériterait d'être approfondie pour avoir une vision globale et non partielle de l'efficacité des politiques de discrimination territoriale aux Etats-Unis.

3 L'impact des exonérations territoriales sur l'activité économique en France

En France, les exonérations de cotisations de sécurité sociale représentaient, toutes mesures confondues, 30 milliards d'euros pour l'année 2010 [ACOSS (2011)]. Même si seulement 4,5% de ce montant global - soit quand même près de 1,5 milliard d'euros annuellement - sont attribuables aux exonérations territoriales, leur effet sur l'activité économique mérite d'être évalué, surtout en cette période de fortes contraintes budgétaires.

Par rapport aux Etats-Unis, le déficit de travaux d'évaluation des politiques de discrimination positive territoriale est patent en France. Il s'explique pour partie par l'ampleur beaucoup plus limitée des dispositifs français et du moindre recul temporel dont on dispose en comparaison aux dispositifs américains, mis en place plus précocement. Pour certaines mesures (ZUS, ZRU, BER, ZRD), il n'existe pas de bilans à proprement parler, les éléments d'évaluation se résumant à des rapports officiels basés sur de simples statistiques descriptives et non sur des approches économétriques. Toutefois, le rythme des études dressant un premier bilan des politiques françaises d'exonérations territoriales s'est accéléré ces dernières années, notamment grâce à la mise à disposition récente de données finement géolocalisées. Dans ce contexte, les travaux concernant les expériences américaines ont offert un incontournable point d'ancrage quant aux méthodes et aux types de données utilisées. Nous allons montrer quelles réponses aux problèmes d'évaluation statistique ont été apportées lors de la transposition à la problématique française.

3.1 Les Zones Franches Urbaines, source d'emplois pour les quartiers en difficultés ?

La politique de la ville, centrée initialement sur l'aménagement urbain et les actions sociales et de sécurité au moment où le "problème des banlieues" émergeait dans les grands ensembles,

s'est progressivement réorientée vers des préoccupations plus économiques. Pour preuve, depuis le début des années 90, et la mise en œuvre des zones urbaines sensibles (ZUS), certains quartiers cumulant de nombreux handicaps (fort chômage, précarité, habitat dégradé, *etc*) bénéficient d'un ensemble d'exonérations sociales et fiscales afin d'encourager les entreprises à s'y implanter. Cette politique territoriale est montée en charge en 1996 avec le "Pacte de relance pour la ville", qui outre l'extension de la liste des ZUS a créé deux autres niveaux de quartiers prioritaires : les *zones de redynamisation urbaine* (ZRU) et les *zones franches urbaines* (ZFU). Les 44 premières ZFU ont été désignées parmi 436 ZRU, elles-même sélectionnées parmi 757 ZUS. Les ZFU ont été choisies parmi les ZRU de plus de 10 000 habitants et ayant de très grandes difficultés socioéconomiques. Ces difficultés étaient mesurées en fonction de critères objectifs tels que le taux de chômage, la proportion de jeunes de moins de 25 ans, la proportion d'individus sans diplôme et le potentiel fiscal par habitant de la commune.

Ce dispositif, qui devait initialement durer 5 ans, a été plusieurs fois étendu et prolongé : 41 nouvelles ZFU - anciennement ZRU - sont entrées en vigueur au 1er janvier 2004 et 15 autres nouvelles en 2006. Les 100 ZFU ainsi créées successivement devaient théoriquement disparaître en décembre 2011, mais le dispositif a été prolongé pour une nouvelle période de cinq ans.

Contrairement aux dispositifs des EZ aux Etats-Unis, celui des ZFU est limité aux petites entreprises (moins de 50 salariés et un chiffre d'affaire inférieur à 10 millions d'euros). Les établissements implantés ou s'implantant dans les ZFU et éligibles bénéficient d'un régime fiscal dérogatoire prévoyant une exonération totale de l'impôt sur les bénéfices et des taxes foncières et professionnelles pendant cinq ans, puis d'une diminution progressive les cinq années suivantes. S'ajoute une exonération des charges sociales patronales totale là-encore pendant les cinq premières années et à un taux décroissant pour une période supplémentaire variant entre 3 et 9 ans selon que l'entreprise emploie plus ou moins cinq salariés. Initialement, tous les emplois étaient exonérés mais sur la fraction des salaires inférieure à 1,5 Smic (1,4 Smic depuis 2006). Depuis 2009, de nouvelles règles d'éligibilité ont été fixées en fonction du salaire horaire, accompagnées d'une certaine dégressivité : les exonérations sont désormais dégressives pour les salaires compris entre 1,4 et 2,4 Smic. En contrepartie, à partir de la troisième embauche bénéficiant de l'exemption, les entreprises doivent respecter une clause d'emploi local qui stipule qu'une partie des nouvelles embauches doivent être réalisées au sein de la zone : cette part était initialement fixée à 20%, elle a été portée à 33% en 2002, puis à 50% en 2012.

Les ZUS et les ZRU proposent des systèmes d'exonérations beaucoup moins généreux. Par exemple, dans les ZRU, les exonérations de cotisations sociales patronales sont uniquement réservées aux nouveaux embauchés dans ces zones et non à l'ensemble des salariés comme dans les ZFU et ceci pour une durée de 1 an et non de 5. De même, les entreprises créant de nouvelles embauches en ZUS peuvent être exemptées du paiement des taxes locales, mais pas de façon systématique. Du fait de ces divers degrés d'incitations, on s'attend à un impact sur l'activité économique *a priori* plus marqué pour les ZFU que les deux autres catégories de zones prioritaires. C'est pourquoi, les économistes se sont concentrés à évaluer les retombées des ZFU. Les ZUS et les ZRU n'ont pas fait l'objet de véritables évaluations économétriques ; mais certains enseignements sur ces zones pourront être tirées des travaux sur les ZFU.

Plusieurs études existent sur l’efficacité des ZFU. Les premières se sont concentrées sur celles de deuxième génération (créées en 2004) car les groupes de contrôle paraissaient plus simples à déterminer [Rathelot & Sillard (2008), Givord, Rathelot & Sillard (2012*b*), Mayer, Mayneris & Py (2012)] alors que les plus récentes, utilisant des méthodes d’évaluation plus originales, se sont intéressées aux ZFU de première génération [Gobillon, Magnac & Selod (2012), Givord, Quantin & Trevien (2012*a*)].

Pour évaluer l’impact des ZFU de deuxième génération sur l’emploi salarié et les créations d’établissements, Rathelot & Sillard (2008), Givord et al. (2012*b*) et Mayer et al. (2012) se sont inspirés des études et des débats concernant les programmes américains de discrimination territoriale. Rathelot & Sillard (2008) et Givord et al. (2012*b*) basent leur évaluation sur un estimateur de différences de différences pondéré [équation (7)] mais en prenant soin d’apporter des réponses aux principales critiques qui avaient été adressées à cette méthodologie. Premièrement, pour déterminer leur groupe de contrôle, ils combinent, comme Busso & Kline (2008), l’approche de caractérisation *ad hoc* et celle d’appariement par score de propension. Pour leur appariement par score de propension, ils ne considèrent pas l’ensemble des zones non ZFU, mais uniquement les ZRU qui ne sont pas passées ZFU (et donc restées ZRU) après 2004. Les arguments en faveur de cette restriction sont similaires à ceux avancés par Moore (2003) et Neumark & Kolko (2010) quand ils considéraient comme meilleur groupe de contrôle pour les EZ initiales les zones qui le devenaient en seconde période. Toutefois, en utilisant les ZRU définies à la vague précédente plutôt que les ZFU de première génération, ils s’affranchissent de la critique selon laquelle prendre deux vagues successives du programme comme groupes de référence risque de biaiser les résultats en raison de la chronologie de sélection basée sur les difficultés des zones. Rathelot & Sillard (2008) et Givord et al. (2012*b*) montrent bien que les ZRU restées ZRU en 2004 et celles qui sont passées ZFU en 2004 sont vraiment très similaires en termes de développement économique car le principal facteur ayant joué dans la désignation des ZFU de deuxième génération était la localisation géographique : les ZRU les plus éloignées d’une ZFU de première génération avaient plus de chances d’être désignées ZFU de deuxième génération. D’ailleurs, pour confirmer la pertinence de ce choix, Givord et al. (2012*b*) ont effectué un test de falsification : ils ne trouvent aucune différence significative en termes d’emplois, de créations d’entreprises, *etc* entre les ZFU de deuxième génération et leur groupe de contrôle avant la mise en place du dispositif d’exonérations. A partir de ces groupes, Rathelot & Sillard (2008) et Givord et al. (2012*b*) appliquent la méthode d’appariement par noyau gaussien sur le score de propension².

Deuxièmement, Rathelot & Sillard (2008) et Givord et al. (2012*b*) sont conscients que l’utilisation de données agrégées au niveau des municipalités pourrait biaiser l’évaluation puisque les contours géographiques des ZFU ne correspondent qu’approximativement à ceux des municipalités. C’est pourquoi, Rathelot & Sillard (2008) se placent à un niveau géographique plus fin : leurs unités de base sont les îlots regroupés pour l’information statistique (IRIS) dont la

2. Les ZRU appartenant au groupe de contrôle sont pondérées selon la distance entre leur score et celui de la ZFU de deuxième génération considérée, les pondérations étant déterminées par un noyau gaussien dont la fenêtre h est fonction de l’écart-type σ du score de propension et du nombre d’observations n selon $h = \sigma.n^{-1/5}$.

taille habituelle correspond à un pâté de maisons. La caractérisation de la variable de traitement à partir de cette échelle géographique n'est pas parfaite mais raisonnable : Rathelot & Sillard (2008) estiment que le nombre d'établissements appartenant à la zone approchée excède d'environ 10% le nombre d'établissements situés dans la zone effective. Pour pallier ce problème d'approximation de la variable de traitement, Givord et al. (2012*b*), comme Mayer et al. (2012), utilisent des données administratives au niveau de l'établissement. Comme ils disposent de leur adresse exacte et d'informations précises sur les contours des zones urbaines géographiquement ciblées, ils construisent une variable de traitement T_i moins sujette aux erreurs de mesure à l'aide d'un système d'information géographique (SIG) et topographique.

Troisièmement, Rathelot & Sillard (2008) et Givord et al. (2012*b*) examinent explicitement les externalités des ZFU sur les quartiers limitrophes non bénéficiaires des politiques d'exonérations. Ils comparent ainsi l'impact moyen du traitement sur les traités avec l'impact moyen sur les zones avoisinantes afin d'avoir une évaluation globale de l'efficacité de ces dispositifs. Enfin, Givord et al. (2012*b*) raffinent et étendent l'étude des ZFU de deuxième génération de Rathelot & Sillard (2008) en élargissant les variables d'intérêt analysées et en examinant l'hétérogénéité du traitement entre secteurs et entreprises (entreprises déjà existantes dans la ZFU avant la désignation versus entreprises nouvellement créées). En particulier, comme Rathelot & Sillard (2008), ils étudient l'impact de ces ZFU sur le nombre d'entreprises et l'emploi total (emplois et heures travaillées), mais ils s'intéressent aussi aux taux de survie des entreprises.

Les résultats de Rathelot & Sillard (2008) et Givord et al. (2012*b*) aboutissent aux mêmes conclusions de faible efficacité économique des ZFU de deuxième génération. La création des 41 ZFU en 2004 aurait eu un impact positif et significatif sur l'activité économique, mais relativement modeste par rapport à son coût. En particulier, ce dispositif aurait favorisé le flux brut d'établissements, avec entre 1 400 et 3 400 établissements supplémentaires estimés fin 2006. Le bilan est d'autant plus mitigé que ces effets seraient entièrement imputables aux nouvelles implantations. Pour les entreprises déjà installées sur ces territoires avant la désignation en ZFU, les importantes exonérations sociales et fiscales n'auraient eu aucun effet significatif sur leur taux de survie, leur niveau d'emploi ou leur santé financière. Ces résultats sont très similaires à ceux trouvés par Bondonio & Greenbaum (2007) et Ham et al. (2011) pour les dispositifs américains malgré la plus grande générosité des incitations pour les ZFU. En outre, les ZFU de deuxième génération semblent avoir créé des effets d'éviction : les deux tiers de l'augmentation du flux d'établissements est attribuable à un déplacement d'activités en provenance de zones non ciblées plutôt qu'à des créations *ex nihilo*. Plus précisément, les zones à proximité immédiate de ces ZFU auraient vu, en 2004, leur nombre d'établissements diminuer de 7%, alors que parallèlement, celui-ci n'aurait augmenté que de 5% dans les ZFU. La plupart des effets directs du dispositif seraient ainsi compensés par des pertes d'activité dans les zones avoisinantes. Par ailleurs, ce programme entraînerait des répercussions hétérogènes selon le secteur d'activité. Le secteur des services aux entreprises serait celui qui aurait le plus bénéficié du programme des ZFU [Givord et al. (2012*b*), Briant et al. (2012)]. Briant et al. (2012) montrent qu'une partie des hétérogénéités serait aussi à relier à des différences d'isolement spatial entre

zones, et en particulier à leurs disparités en termes de connections aux réseaux de transports urbains ou d'accessibilité.

Mayer et al. (2012) affinent ces résultats en examinant l'impact des ZFU de deuxième génération au niveau intra-municipal pour les municipalités bénéficiaires du dispositif. Ils pensent que les effets de transferts d'activités ou d'établissements évoqués par Rathelot & Sillard (2008) et Givord et al. (2012b) pourraient majoritairement avoir lieu entre quartiers d'une même municipalité. C'est pourquoi, ils étudient une variable d'intérêt originale : la probabilité qu'un établissement décide de s'implanter sur le territoire ZFU de la municipalité bénéficiaire plutôt que dans les quartiers non-ciblés de la même municipalité. Ils utilisent un estimateur de différences de différences qui se base sur la variation de cette probabilité avant et après la mise en place du dispositif de ZFU de deuxième génération. Cet estimateur est néanmoins calculé à partir de régressions logit du fait de la nature dichotomique de leur variable d'intérêt. Pour s'assurer de la robustesse de leurs résultats, ils mettent aussi en œuvre des estimations en triple différences en utilisant comme groupe de contrôle les ZFU de troisième génération et ils exploitent la discontinuité à 50 salariés et à 10 millions d'euros du chiffre d'affaires dans les règles d'éligibilité pour proposer des tests de falsification. En effet, l'utilisation de la discontinuité permet d'obtenir des résultats qui ne sont entachés ni du biais de sélection sur les inobservables ni du biais de simultanéité grâce à la source de variation exogène (les seuils imposés) dans le traitement. La validité interne de leurs résultats est ainsi vérifiée puisque les trois méthodologies proposées conduisent aux mêmes conclusions. Plus précisément, Mayer et al. (2012) montrent que le dispositif de ZFU de deuxième génération a eu un impact significatif et positif sur les décisions de localisation des établissements. La probabilité pour un établissement de se localiser dans les territoires ciblés de la municipalité plutôt que dans les quartiers non-ZFU a augmenté en moyenne de 2,6% une fois le dispositif d'exonérations réellement mis en place (2004-2007). Comme la probabilité moyenne de s'implanter dans la partie ZFU d'une municipalité était de 8,4% sur la période 2000-2003, l'impact marginal estimé correspond à une élasticité de 31%. Toutefois, derrière cet impact positif moyen se cachent d'importantes disparités. Le dispositif s'est révélé plus efficace dans les territoires ciblés les moins défavorisés et pour les établissements des secteurs où les coûts de relocalisation étaient les plus faibles (santé, services aux entreprises, commerce). Surtout, cette étude confirme que le dispositif n'a pas créé d'activité économique *per se* au niveau de la municipalité ; il a seulement généré de forts effets de transferts au niveau intra-municipal. L'effet moyen est principalement attribuable à des (re)localisations opportunistes d'établissements entre quartiers ZFU et non-ZFU d'une même municipalité. Le coût total du dispositif paraît donc bien élevé pour des bénéfices si minces.

Givord et al. (2012a) et Gobillon et al. (2012) ont complété ce bilan sur les ZFU en examinant celles de première génération et en apportant quelques nouveautés dans les méthodes d'évaluation utilisées. L'étude de Givord et al. (2012a) s'inscrit dans la lignée de celle de Givord et al. (2012b). D'une part, elle s'intéresse aussi à l'impact des exonérations territoriales sur les stocks d'entreprises et le niveau d'emploi en distinguant les effets liés à de réelles créations de ceux liés à des transferts d'activités entre zones. D'autre part, elle utilise une méthodologie

similaire basée sur un estimateur de différences de différences où le groupe de contrôle est déterminé par appariement sur score de propension à partir des seules ZRU³. Elle présente néanmoins deux originalités. Premièrement, elle utilise des données administratives pour la période 1995-2007 et peut ainsi analyser l’impact des ZFU créées en 1997 sur les 10 années après leur mise en place. Deuxièmement, Givord et al. (2012a) préfèrent recourir à une *méthode de différences de différences avec une classification sur le score de propension (subclassification or stratification on the propensity score)* plutôt que d’utiliser ce dernier pour un appariement standard par noyau. Cette méthode, suggérée initialement par Rosenbaum & Rubin (1983), consiste à diviser l’échantillon en sous-classes, basées sur la valeur du score de propension, de telle sorte qu’au sein de chaque sous-classe le score de propension soit approximativement constant. L’avantage de cette classification est que l’on peut estimer l’effet moyen du traitement au sein de chaque sous-classe comme si les données étaient issues d’une expérimentation où l’assignation entre le groupe des traités et le groupe de contrôle avait été complètement aléatoire. En effet, les biais liés à l’hétérogénéité potentielle entre les deux groupes⁴ peuvent être réduits par la classification sur scores de propension puisqu’elle forme des classes où les zones des deux groupes ont des scores de propension très proches et donc des distributions de caractéristiques socioéconomiques vraiment très comparables. Pour estimer l’effet moyen global du traitement sur les traités Δ , il suffit alors de faire la moyenne des effets moyens du traitement estimés sur chaque sous-classe, pondérée par la taille de la sous-classe.

Plus précisément, Givord et al. (2012a) déterminent 4 sous-classes à partir des valeurs du score de propension et sur chacune d’elles effectuent les régressions linéaires suivantes :

$$\ln(Y_{it}) - \ln(Y_{it'}) = X_i\beta_j + \delta_j T_i + u_{ij}, j = 1, \dots, 4 \quad (10)$$

où X_i correspond aux caractéristiques socioéconomiques de la zone i et où δ_j reflète l’effet du traitement sur les traités pour la sous-classe j .

L’effet moyen du traitement est alors estimé selon la moyenne pondérée suivante :

$$\Delta = \sum_{j=1}^4 \hat{\delta}_j \cdot (N_j/N) \quad (11)$$

où N_j est le nombre d’observations dans la sous-classe j et N le nombre total d’observations (les ZFU de première génération et les ZRU avec les caractéristiques les plus proches, servant de groupe de contrôle).

Givord et al. (2012a) mettent en évidence un fort effet positif des ZFU de première génération sur l’emploi et les implantations d’entreprises pendant les premières années du dispositif. Au bout de 5 ans, les ZFU créées en 1997 auraient favorisé l’installation de 9 700 à 12 200 établissements supplémentaires, soit entre 41 500 et 56 900 emplois. Cet effet dépasse de

3. Ils estiment aussi l’impact des ZFU avec l’approche par discontinuité de la régression (cf. section 3.2), mais plus pour contrôler la robustesse de leurs résultats.

4. Rosenbaum & Rubin (1984) parlent du “covariate imbalance” bias.

loin les chiffres des évaluations des ZFU de deuxième génération. Rappelons à titre de comparaison que ces dernières n'auraient permis l'implantation que de 1 400 à 3 400 établissements supplémentaires sur le même laps de temps [Givord et al. (2012b)]. Toutefois, les effets des ZFU de première génération semblent avoir plafonné à moyen terme puisque le nombre d'établissements et d'emplois aurait stagné à partir de 2002. Cet essoufflement s'explique en partie par l'augmentation progressive des disparitions d'établissements dans ces zones, à partir de 2002, c'est-à-dire juste à la fin de la période ouvrant droit à une exonération totale des taxes sociales et fiscales. Ainsi, au bout de 5 ans de dispositif, l'augmentation des créations d'entreprises était complètement contrebalancée par l'accélération des fermetures d'établissements. En outre, cette étude confirme que le dispositif d'allègements fiscaux des ZFU n'aurait eu aucune influence significative sur l'emploi des entreprises déjà implantées dans ces zones et que la majorité des nouvelles implantations seraient dues à des relocalisations d'activités initialement existantes hors ZFU plutôt qu'à la création effective de nouveaux établissements. Enfin, l'effet positif des ZFU de première génération serait essentiellement imputable à des établissements de services aux entreprises dont on peut craindre qu'elles aient peu d'impact sur l'emploi local ou l'activité économique locale puisqu'elles n'ont pas besoin notamment d'opérer sur sites.

D'ailleurs, Gobillon et al. (2012) confirment que les effets positifs des ZFU de première génération sur les créations d'entreprises ne se sont pas traduits par une forte hausse de l'accès à l'emploi pour la population de ces zones ciblées. Ils estiment sur des données de l'Ile-de-France une régression sur score de propension très proche de l'équation (9), mais en présentant une double originalité. Premièrement, leur problématique est d'évaluer l'impact des ZFU non pas sur la croissance de l'emploi ou les créations d'entreprises dans ces territoires comme la majorité des travaux précédents, mais sur le retour à l'emploi des chômeurs résidant dans ces zones. Le but de leur étude est de savoir si la création des ZFU, implicitement en attirant de nouvelles entreprises, en rapprochant les emplois de cette population et leur accessibilité, a bien eu pour conséquence la réduction de la durée de chômage de ses habitants, objectif affiché lors de la mise en place de ces dispositifs. Ainsi, leur variable d'intérêt dans l'équation (9) n'est pas le nombre d'emplois créés, mais une variable reflétant les *effets de municipalités* qui décrivent la facilité (via la durée) d'un chômeur à trouver un emploi dans la municipalité i à la période t . Deuxièmement, dans la pratique, leur stratégie d'estimation de la régression sur score de propension est beaucoup plus complexe que celle de Bondonio & Engberg (2000) ou Bondonio & Greenbaum (2007), car elle requiert en amont deux estimations auxiliaires : i) bien évidemment une régression probit expliquant la probabilité que la municipalité i soit choisie pour accueillir une ZFU ou non, et qui va permettre d'estimer les scores de propensions $\hat{P}R_i^j$, introduits ensuite dans l'équation (9) ; ii) un modèle de durée afin de déterminer les taux de sorties du chômage vers l'emploi dans les différentes municipalités à chaque période de temps, purgés des effets de caractéristiques socio-démographiques (sexe, âge, diplôme, handicap, expérience, etc). Ces taux de sorties, appelés *effets de municipalités*, sont alors utilisés pour construire la variable dépendante en différence de la régression sur score de propension. Toutefois, comme les travaux précédents, Gobillon et al. (2012) restreignent leur échantillon aux municipalités contenant une ZFU (groupe des traités) et des municipalités qui leur sont proches en termes de caractéristiques et de localisation. Le groupe de contrôle est en effet constitué à partir de

critère de taille - les municipalités avec des ZFU sont généralement de plus grande taille et donc sont enlevées du groupe de contrôle les trop petites municipalités - et d'une restriction sur le score de propension - pour appartenir au groupe de contrôle, il faut avoir un score supérieur au 10ème percentile du score de propension du groupe des traités.

A partir d'une base de données d'individus inscrits au chômage en région parisienne et couvrant la période 1993-2003, Gobillon et al. (2012) concluent que la première génération de ZFU a eu un effet positif mais très faible : les chômeurs dans les municipalités où a été créée une ZFU ont des chances légèrement plus élevées (environ 3% de plus) de retrouver un emploi que les chômeurs dans les municipalités similaires qui n'ont pas bénéficié d'un tel dispositif. En outre, ces effets restent très localisés et ne sont significatifs qu'à court terme. Les résultats de Gobillon et al. (2012) confortent l'idée qu'il fallait peut-être renforcer la clause d'emploi local de ces dispositifs, mais cela risque de freiner les embauches globales.

L'ensemble des évaluations des ZFU montre que les entreprises seraient bien sensibles aux exonérations de charges sociales et fiscales. Mais leurs résultats remettent en doute la capacité de tels dispositifs d'améliorer durablement l'économie locale des zones territoriales ciblées. Les effets positifs sur l'emploi sont relativement modestes, transitoires et ils reflètent essentiellement des effets d'aubaine et d'évictions. Par ailleurs, l'écart observé en termes d'efficacité économique entre la première et la deuxième génération de ZFU nous permet d'avoir des doutes sur la pertinence de la prolongation de ces dispositifs parallèlement au maintien de l'important système d'exonérations de charges sur les bas salaires. En effet, le développement des allègements de charges sur les bas salaires à l'échelle nationale, via notamment la loi "Fillon" en 2003, a réduit l'avantage comparatif des ZFU. Si les exonérations des ZFU sont restées intéressantes pour les salaires élevés, l'avantage est devenu quasiment nul au niveau du Smic. L'écart de taux de cotisations entre ZFU et hors ZFU est passé de 10 points en 1997 à 2 points en 2006, au niveau du Smic [Givord et al. (2012a)]. Les ZRU et les ZUS, aux exonérations encore moins généreuses, ne peuvent dans ce contexte jouer un rôle majeur dans la redynamisation de l'activité économique de ces territoires. On peut craindre aussi un bilan très modeste des ZFU de troisième génération, surtout en comparaison aux coûts qui y sont associés.

3.2 Les zones rurales plus faciles à redynamiser que les zones urbaines ?

Dans une logique similaire à la politique de la ville pour la redynamisation économique des quartiers défavorisés urbains, la France a défini un système de zonage à deux niveaux pour les zones rurales. Les *territoires ruraux de développement prioritaire* (TRDP) sont constitués en 1989, et parmi ces derniers, les *zones de revitalisation rurale* (ZRR) mises en place en 1996 regroupent les territoires les plus fragiles. De nouveau, les ZRR bénéficient d'incitations fiscales dans le but affiché d'attirer de nouvelles entreprises et d'amorcer un regain d'activité, qui devait ensuite être soutenu par l'arrivée de nouveaux ménages dans ces zones.

La nature du dispositif des ZRR est tout à fait comparable à celui des ZFU, même s'il diffère par son ampleur. Premièrement, à leur création, les allègements en ZRR ne concernaient

que les cotisations sociales patronales, alors qu'en ZFU s'ajoutent des exonérations de taxes professionnelles, de taxes foncières et d'impôt sur les bénéfices. Deuxièmement, l'exonération de cotisations sociales patronales est moins généreuse en ZRR. Certes, comme en ZFU, elle est totale pour la part des salaires inférieure à 1,5 Smic, mais elle ne concerne que les nouvelles embauches (et non l'ensemble des salariés comme en ZFU) et est limitée à une durée de 1 an (contre 5 ans pour les ZFU). Enfin, cette exonération est assujettie à un nombre de conditions relativement restrictives réduisant d'autant le nombre de bénéficiaires potentiels : elle s'applique pour les entreprises de moins de 50 salariés, dont l'effectif salarié est en augmentation et qui n'ont réalisé aucun licenciement dans les deux dernières années. En outre, pour être éligibles, les entreprises en ZRR devaient remplir les critères démographiques et économiques suivants : être incluse dans un canton dont la densité de population est inférieure ou égale à 31 habitants/km² ou un arrondissement de densité inférieure à 33 habitants/km² ; être dans une commune ayant connu une diminution de sa population ou de sa population active supérieure à deux fois la moyenne nationale ou un niveau d'emploi agricole 2 fois supérieur à la moyenne nationale.

Or, le faible taux de souscription au dispositif des ZRR parmi les établissements éligibles, quelques années après son instauration, a résonné comme un premier constat d'échec. En 2000, seulement 3,7% des 190 000 établissements de moins de 50 salariés situés en ZRR ont bénéficié d'exonérations [Lorenceau (2009)]. C'est pourquoi, le dispositif des ZRR a été profondément modifié en 2005. En termes des allègements fiscaux concernés (taxe professionnelle, taxes foncières, impôt sur les bénéfices) et des durées des exonérations (de 2 à 5 ans), il est désormais très proche de celui des ZFU.

La première véritable évaluation de l'impact économique des ZRR est celle de Lorenceau (2009) qui a étudié les retombées du dispositif initial d'allègements en s'intéressant à l'évolution de l'emploi et des implantations d'entreprises dans ces zones sur la période 1995-1998. Behaghel, Lorenceau & Quantin (2012) l'ont complété récemment en tenant compte des critiques principales qui lui avaient été adressées et en étendant l'étude au système plus généreux mis en place en 2005. Malgré les similitudes des dispositifs des ZRR et des ZFU, Lorenceau (2009) et Behaghel et al. (2012) se démarquent, par rapport aux travaux sur les zones urbaines, de par leur méthodologie d'évaluation statistique. Ils privilégient l'approche par *régression sur les discontinuités* (appelée aussi *discontinuité de la régression*, *regression discontinuity design*, *RDD*) aux méthodes de différences de différences avec ou sans appariement. Cette préférence pour les régressions sur discontinuités réside tout d'abord dans le fait qu'il semblait difficile au vu des données de trouver des ZRR et des non-ZRR aux caractéristiques similaires : les cantons qui bénéficiaient des exonérations ne ressemblaient pas à ceux qui avaient été exclus du dispositif. En outre, contrairement aux méthodes d'appariement par score de propension, l'approche par régressions sur discontinuités s'affranchit de l'*hypothèse d'indépendance conditionnelle* - qui postule qu'il n'existe pas de variables non prises en compte qui interviennent à la fois dans le processus de sélection des zones traitées et dans les performances économiques de celles-ci - loin d'être toujours garantie et difficilement testable. Toutefois, cela se fait au prix d'autres hypothèses stochastiques.

Les régressions sur discontinuités exploitent de manière intuitive l’existence de seuils dans la détermination des territoires ciblés. L’idée est qu’autour du seuil imposé, des zones très proches peuvent avoir des chances très différentes de bénéficier du dispositif. En l’occurrence pour les ZRR, Lorenceau (2009) et Behaghel et al. (2012) utilisent le seuil de densité de 31 hab/km². Il est certainement faux de supposer que les cantons ZRR d’une densité inférieure à ce seuil auraient connu, en l’absence du dispositif ZRR, le même développement économique en moyenne que les cantons non-ZRR d’une densité supérieure ; en revanche, cette hypothèse devient acceptable dès que l’on restreint la comparaison aux cantons ZRR juste en dessous du seuil et ceux non-ZRR juste au-dessus. Il est donc possible d’identifier l’effet causal du traitement en comparant ces deux groupes de zones. La méthode de régressions sur discontinuités n’exploite ainsi qu’une partie de la source de variabilité dans l’affectation du statut ZRR, qui peut être raisonnablement supposée exogène. Pour pouvoir mettre en œuvre cette méthode, l’attribution du traitement doit avoir été déterminée, au moins en partie, par une dimension quantifiable et observée, appelée généralement *variable d’affectation*, en l’occurrence la densité démographique dans le cas des ZRR. En d’autres termes, les régressions sur discontinuités reposent sur l’existence d’une variable d’affectation (S) qui a un impact discontinu sur la probabilité d’être traitée. Comme la désignation en ZRR ne s’est pas uniquement basée sur la seule valeur de la densité démographique des zones, on est dans le cadre d’un *fuzzy design*. L’effet du traitement sur les traités est alors mesuré par ratio entre l’ampleur de la discontinuité dans la variable d’intérêt et l’ampleur de la discontinuité dans la probabilité d’être traitée [Hahn, Todd & van der Klaauw (2001)] :

$$\Delta = \frac{\lim_{S \searrow \underline{S}^+} E(Y|S) - \lim_{S \nearrow \underline{S}^-} E(Y|S)}{\lim_{S \searrow \underline{S}^+} E(T|S) - \lim_{S \nearrow \underline{S}^-} E(T|S)} \quad (12)$$

avec S la variable d’affectation (densité de la population), \underline{S} le seuil d’inclusion (31 hab/km²), \underline{S}^+ la limite à droite de \underline{S} et \underline{S}^- la limite à gauche.

La condition d’identification pour que cet effet soit correctement estimé est que la variable d’intérêt (création d’emplois et d’entreprises) est une fonction continue et lisse de la variable d’affectation (densité de la population), notamment autour du seuil (31 hab/km²). Or, pour que cette hypothèse soit valide, il faut que le seuil choisi n’ait pas pu être “manipulé” en fonction des résultats escomptés. Ainsi, dans notre contexte des ZRR, il faut se demander si le seuil de 31 hab/km² a réellement été déterminé de façon arbitraire par rapport aux variables d’intérêt ou en d’autres termes si la valeur de ce seuil n’est pas le résultat de pressions, négociations avec des élus locaux pour rendre leurs cantons éligibles à la mesure. Lorenceau (2009) et Behaghel et al. (2012) affirment que la décision concernant la valeur du seuil a été prise au niveau ministériel sur des critères purement budgétaires et non dans l’optique d’inclure ou d’exclure telles ou telles zones particulières. C’est pourquoi, ils considèrent que le fait qu’un canton se trouve juste au dessus ou juste en dessous du point seuil n’est lié *a priori* ni à l’emploi, ni à la création d’entreprises et ni de manière générale aux bénéfices qu’il peut attendre de la mise en

œuvre de la politique. L’hypothèse d’identification leur paraît donc vérifiée, ce qui est confirmé par certains “tests graphiques” [voir Givord (2010) pour plus de détails sur ces tests].

Dans la pratique, on peut obtenir l’estimateur de l’équation (12) à partir d’une régression linéaire sur la variable d’intérêt (Y_i) après avoir instrumenté le fait d’avoir bénéficié du traitement T_i et en restreignant l’échantillon d’estimation aux zones au voisinage du seuil, c’est-à-dire appartenant à l’intervalle $[\underline{S} - h, \underline{S} + h]$:

$$Y_i = \alpha + f(S_i, \underline{S}) + \delta \hat{T}_i + \varepsilon_i \quad (13)$$

où f est une fonction flexible de la variable d’affectation et du seuil.

La variable T_i est usuellement instrumentée par la variable indicatrice $\mathbf{1}_{\underline{S}}$ qui vaut 1 si S est inférieur à \underline{S} et 0 sinon. Pour leur application aux ZRR, Lorenceau (2009) et Behaghel et al. (2012) utilisent pour la fonction f une forme linéaire et un polynôme du second degré de la variable S et de son interaction avec la variable $\mathbf{1}_{\underline{S}}$. Dans cette équation, δ s’interprète comme le supplément de croissance de l’emploi ou du nombre d’entreprises que l’on observe pour les cantons classés en ZRR du fait de la mise en place du dispositif.

Au-delà des hypothèses sous-jacentes, le point délicat dans l’utilisation des régressions sur discontinuités est le choix de la largeur de la bande h restreignant l’échantillon d’estimation. Si h est choisi trop petit, on prend le risque d’avoir une estimation imprécise du fait du faible nombre de points utilisés⁵ et si h est trop grand, un biais est susceptible d’apparaître puisque les zones éloignées de la discontinuité ont moins de chances d’avoir des caractéristiques similaires. En outre, une des limites des régressions sur discontinuités est qu’elles estiment seulement *un effet moyen local du traitement*, puisqu’il est uniquement évalué aux alentours du seuil d’inclusion. Ainsi, les résultats ne sont pas toujours généralisables et on retombe sur un *problème de validité externe*.

En appliquant cette méthodologie avec précaution, Lorenceau (2009) conclut que le dispositif initial d’exonérations fiscales à l’embauche en ZRR n’a eu aucun impact significatif sur les taux de croissance de l’emploi et du nombre d’établissements dans ces zones ciblées entre 1995 et 1998. Toutefois, Riedinger (2009) dans un commentaire de cette étude souligne plusieurs faiblesses. Premièrement, les données utilisées par Lorenceau (2009) sont agrégées au niveau du canton alors que l’exploitation directe de données individuelles d’établissements aurait permis une estimation plus précise. Deuxièmement, Riedinger (2009) regrette que Lorenceau (2009) ait limité son étude à la période 1995-1998 puisque le dispositif ZRR n’a atteint son plein régime qu’en 1997. Ainsi, la période d’évaluation est trop courte pour entrevoir l’impact à moyen et à long terme. Enfin, Riedinger (2009) affirme qu’il ne faut pas confondre effet “statistiquement significatif” et effet “économiquement significatif”. Ainsi, le fait que l’effet des ZRR estimé ne

5. Givord et al. (2012a) réalisent aussi une approche par régressions sur discontinuités pour évaluer l’impact des ZFU de première génération. Néanmoins, ils avancent que leur estimation par la méthode de différences de différences avec classification sur le score de propension est plus robuste du fait du peu d’observations dont ils disposent dans l’intervalle restreint autour du seuil d’inclusion.

diffère pas significativement de 0 ne signifie pas forcément que le dispositif n'est pas efficace. Cela pourrait seulement suggérer que les régressions sur discontinuités n'arrivent pas à mesurer l'impact moyen avec assez de précision. C'est pourquoi, Behaghel et al. (2012) ont récemment réévalué le dispositif des ZRR en tenant compte de l'ensemble de ces critiques. Ils utilisent des données administratives au niveau de l'établissement plus fiables et leur évaluation couvre une période plus longue : 1994-2004 pour le dispositif initial et 2005-2008 pour les nouvelles dispositions plus généreuses.

Behaghel et al. (2012) confirment pourtant que le dispositif initial des ZRR n'aurait eu aucun effet significatif sur la croissance de l'emploi et les créations nettes d'entreprises de ces zones, à court terme et à long terme. Par ailleurs, ils trouvent que le dispositif plus généreux mis en place en 2005 n'aurait malheureusement pas eu plus de retombées positives. Or, ils montrent que cette absence d'impact n'est pas liée à un manque de puissance statistique de leurs régressions mais à une réelle inefficacité économique du dispositif. Ces résultats pour les ZRR renforcent les conclusions obtenues pour les ZFU de relative faible efficacité économique des politiques d'exonérations territoriales. Il n'est pas surprenant que les retombées économiques aient été encore plus faibles que dans le cas des ZFU puisque les exonérations proposées par le dispositif des ZRR étaient moins généreuses, uniquement pour les nouvelles embauches et de très courte durée. Le dispositif était donc peu attractif par rapport aux allègements de charges mis en place à l'échelle nationale sur les bas salaires, qui concernaient tous les salariés avec un salaire inférieur à 1,3 Smic et qui n'avaient pas une durée limitée. La modification du dispositif des ZRR en 2005 pour le rapprocher de celui des ZFU et le rendre plus attractif ne semble pas avoir amélioré son efficacité économique. Il se pourrait donc que ce dispositif soit supprimé prochainement comme certains politiques le souhaiteraient en cette période de restrictions budgétaires. De même, on peut être relativement sceptique sur l'efficacité des dispositifs d'exonérations relatifs aux *bassins d'emplois à redynamiser* (BER) et aux *zones de restructuration de la défense* (ZRD), mis plus récemment en place mais sur des bases relativement similaires aux ZRR. Dans ces deux derniers cas, comme les variables d'affectation sont moins évidentes que dans le cas des ZRR, des méthodes d'évaluation de différences de différences ou de régressions sur les scores de propension seraient *a priori* à privilégier.

3.3 Des politiques d'exonérations efficaces pour les territoires ultramarins ?

S'élevant à 917 millions d'euros, les exonérations en faveur des entreprises des départements d'Outre-mer totalisaient presque les trois-quarts de l'ensemble des exonérations zonées en 2010 [ACOSS (2011)]. Par ailleurs, ces territoires ultramarins bénéficient d'un régime d'exonérations de cotisations sociales patronales très spécifique, depuis sa mise en place par la loi "Perben" en 1994 et ses modifications successives en 2001 (loi d'orientation pour l'Outre-mer, "Loom"), en 2003 (loi de programmation pour l'Outre-mer, "Lopom") et en 2009 (loi de programme pour le développement économique et la promotion de l'excellence Outre-mer, "Lodéom"). La particularité de leur dispositif d'exonérations territoriales s'inscrit dans la lignée de leur statut de "région ultrapériphérique" au sein de l'Union Européenne qui leur permet de déroger au

droit commun métropolitain et européen afin de compenser les coûts économiques et sociaux de l'éloignement de la métropole et de l'insularité. Par exemple, en 2003, les exonérations de charges patronales sont ciblées sur des secteurs prioritaires dans les DOM et différenciées selon leur exposition à la concurrence internationale et la taille des entreprises⁶. De plus, le montant de l'allègement augmentait avec le salaire jusqu'à un certain seuil. Certes, la loi "Lodéom" de 2009 a atténué les spécificités du régime ultramarin, notamment en unifiant le plafond des exonérations (à 1,4 Smic) et en introduisant une dégressivité des réductions, mais elle a maintenu un système d'exonérations renforcé pour les entreprises de certains secteurs d'activité ayant moins de 250 salariés.

Pourtant, malgré la grande ampleur de ce dispositif et ses particularités, aucune évaluation économétrique de ses retombées en termes d'emplois n'a été effectuée. Comme le note le rapport de l'Assemblée Nationale relatif au projet de loi de finances 2011, « le principal obstacle à une évaluation de l'impact de ces exonérations réside dans la difficulté à isoler les seuls effets des exonérations, l'Outre-mer bénéficiant d'un ensemble de dispositifs spécifiques censés favoriser le développement économique, par exemple la défiscalisation des investissements ou la taxe sur la valeur ajoutée non perçue récupérable ». En outre, peu de données statistiques fiables sont mobilisables dans le cas domien pour mener une évaluation *ex-post* comparable à celles des dispositifs de ZFU ou ZRR.

En conséquence, les politiques d'exonérations pour les territoires ultramarins souffrent actuellement d'un manque patent de véritables évaluations statistiques de leur efficacité. Comme aides à la décision sur cette problématique, les pouvoirs publics doivent se contenter pour l'instant de l'étude de Baudin, Legendre & L'Horty (2010, 2011) qui *simule les effets* de la politique de discrimination territoriale à partir d'une maquette théorique calibrée. La robustesse des résultats simulés et donc *ex-ante* est bien moindre que celle des estimations microéconométriques *ex-post*. C'est pourquoi, les prédictions de Baudin et al. (2010, 2011) mériteraient d'être réfutées ou confirmées empiriquement.

La simulation des effets est basée sur des hypothèses très fortes et souvent *ad hoc* tant au niveau de la spécification du modèle théorique utilisé que du calibrage de celui-ci. Or, les résultats qualitatifs et quantitatifs obtenus sont vraiment conditionnels à la calibration choisie des paramètres du modèle et en particulier des élasticités. Même s'il ne faut pas lui donner le même crédit que les évaluations des autres politiques géographiquement ciblées, intéressons-nous à l'étude de Baudin et al. (2010, 2011).

Baudin et al. (2010) examinent le dispositif des exonérations de cotisations sociales patronales qui a été en vigueur dans les DOM de 2003 à 2009. Pour appréhender aux mieux ses effets, ils se basent sur une représentation théorique multisectorielle et multicatégorielle d'une petite économie ouverte. Chaque secteur d'activité se distingue selon son degré d'ex-

6. L'exonération des charges sociales patronales est totale jusqu'à un seuil de 1,5 Smic pour les secteurs les plus exposés à la concurrence internationale, essentiellement le tourisme, jusqu'à un seuil de 1,4 Smic pour les secteurs qui relèvent principalement de l'industrie et de l'agriculture et jusqu'à un seuil de 1,3 Smic pour les secteurs du transport et du bâtiment et toutes les entreprises de moins de 10 salariés.

position à la concurrence internationale et selon l'intensité de ses relations de concurrence avec l'économie informelle. Neuf catégories de main d'œuvre sont distinguées car la législation de 2003 différencie les montants des allègements offerts selon le niveau de rémunération, ce qui peut générer des phénomènes de réorganisation de la production entre catégories. Plus précisément, Baudin et al. (2010) considèrent que les exonérations de cotisations sociales dans les DOM ont une répercussion sur l'emploi via deux principaux canaux de transmission : *l'effet de compétitivité* et *l'effet de substitution* entre les catégories de main d'œuvre. D'une part, une exonération de cotisations sociales dans un secteur particulier augmente la compétitivité prix de ce secteur, ce qui lui fait gagner des parts de marché à la fois sur ses concurrents dans le reste du monde et sur l'économie informelle. Ainsi, cet effet de compétitivité a un impact positif sur l'emploi de toutes les catégories de main d'œuvre de ce secteur. D'autre part, la différenciation des exonérations selon les rémunérations suscite des réorganisations de production entre les différentes catégories de main d'œuvre, certaines voyant leur coût plus diminuer que d'autres. Cet effet de substitution dépend de la variation relative des coûts d'usage entre chaque catégorie et de leurs possibilités techniques de substitution. A priori, ces possibilités de substitution sont d'autant plus élevées que les catégories sont "proches". Or cette réorganisation de la production peut entraîner une baisse de l'emploi pour une catégorie de la main d'œuvre même si le coût de celle-ci diminue. En conséquence, l'impact total de l'effet de compétitivité et de l'effet substitution va dépendre des élasticités de la demande de travail et de la demande de biens considérés, différentes pour chaque secteur. En outre, dans leur modèle, Baudin et al. (2010) tiennent compte d'une particularité de l'économie domienne en introduisant des *effets d'entraînement* hiérarchisés. Ils supposent que le supplément d'activité engendré par les dispositifs d'exonérations dans les secteurs les plus exposés à la concurrence se répercute positivement sur l'emploi des secteurs abrités, alors que ce n'est pas le cas en sens inverse. Les secteurs les plus abrités bénéficient ainsi à plein des effets d'entraînement puisqu'ils profitent des suppléments d'activités générés par l'ensemble des secteurs plus exposés qu'eux.

Après calibrage de leur maquette à partir des déclarations annuelles de données sociales (DADS) et de sources administratives départementales, Baudin et al. (2010) obtiennent, pour chaque DOM et chaque secteur, l'impact d'une baisse des cotisations sur l'emploi total. Leurs simulations concluent que près de 10% des emplois auraient été supprimés si les DOM n'avaient pas bénéficié du dispositif spécifique de la loi Lopom de 2003. Le secteur le plus touché par ces pertes d'emplois aurait été celui du tourisme où un emploi sur cinq aurait été perdu. Mais cela aurait aussi eu un impact non négligeable sur les secteurs non prioritaires, qui ne bénéficiaient pas des allègements : du fait des effets d'entraînement, l'emploi aurait diminué d'environ 4% dans ces secteurs non éligibles aux exonérations.

L'évaluation ex-ante de Baudin et al. (2010) suggère ainsi que les exonérations territoriales dans les DOM auraient eu des retombées positives sur l'activité économique entre 2003 et 2009 et que la différenciation des allègements de charges selon les secteurs d'activité serait pertinente dans le contexte domien. Toutefois, Baudin, Legendre & L'Horty (2011), avec la même maquette, mettent en évidence que l'efficacité du dispositif ultramarin d'exonérations des cotisations sociales patronales aurait pu, à dépenses constantes, être accrue avec la mise en

place d'un autre barème, ciblant davantage les réductions sur les travailleurs peu qualifiés et sur les secteurs les plus exposés à la concurrence internationale, qui font bénéficier de suppléments d'activités aux secteurs les plus abrités par effet de déversement. En outre, selon Baudin et al. (2010), la réforme du dispositif entrée en vigueur en 2009, notamment pour des raisons budgétaires, devrait engendrer une baisse de l'emploi d'environ 1,4% dans l'ensemble des DOM.

En résumé, pour la France métropolitaine, les évaluations existantes sur l'impact des ZFU et des ZRR trouvent des résultats cohérents entre elles et dans la lignée des travaux sur les expériences américaines. Ces dispositifs d'exonérations territoriales ont des effets incertains [Lorenceau (2009), Behaghel et al. (2012)] et relevant davantage de déplacements d'emplois que de réelles créations nettes d'emplois [Rathelot & Sillard (2008), Givord et al. (2012*b*), Mayer et al. (2012), Givord et al. (2012*a*)]. Par ailleurs, les chômeurs de ces zones ciblées semblent peu profiter de l'activité supplémentaire générée [Gobillon et al. (2012)], ce qui était pourtant l'un des objectifs affichés de ces politiques de discrimination positive. Les simulations de Baudin et al. (2010) laissent penser que les gains sur l'activité locale pourraient être plus perceptibles pour les dispositifs d'Outre-mer, qui sont à la fois les plus coûteux et les plus généreux. Toutefois, ces résultats mériteraient d'être confirmés par une véritable évaluation économétrique sur données *ex-post*.

L'efficacité des dispositifs d'exonérations territoriales en France a ainsi été peu prouvée empiriquement. La plupart des études justifient leurs faibles retombées économiques par leur manque d'attractivité par rapport à la politique, à l'échelle nationale, d'allègements de charges sur les bas salaires. Par exemple, le dispositif pour les ZFU n'a pas été modifié à la suite de la mise en place de la loi "Fillon" alors que celle-ci a considérablement modifié sa portée. Du fait de la forme de l'allègement général et de son intensité au niveau du Smic, les exonérations géographiquement ciblées ne présentent qu'un avantage réduit par rapport à l'allègement Fillon au niveau du Smic [Amghar & Laloue (2010)]. En outre, leur avantage différentiel tend plutôt à croître avec le salaire pour atteindre un maximum au seuil plafond des exonérations (1,4 Smic ou 1,6 Smic voire au delà pour certains dispositifs dans les DOM). Par conséquent, l'existence de l'allègement général conduit à accorder un niveau d'aide maximal à des niveaux de salaires relativement élevés, qui correspondent peu aux qualifications de la main d'œuvre initialement visée de ces zones par la discrimination positive. Ces mécanismes augmentent ainsi les risques d'effets d'aubaine. C'est pourquoi, Amghar & Laloue (2010) plaident pour un dispositif unique d'exonérations territoriales en métropole, qui soit davantage concentré sur les niveaux de salaires les plus bas et dégressif.

4 Conclusion

Cet article recense la principale littérature concernant l'évaluation des impacts sur le marché du travail des politiques de discrimination positive territoriale aux Etats-Unis et en France. Il en ressort que les résultats de telles politiques sont mitigés en termes de créations nettes d'emplois, de créations d'entreprises et de retour à l'emploi des chômeurs des zones ciblées.

Pour la France, l'efficience de ces politiques d'exonérations territoriales pourrait sans doute être améliorée en harmonisant la plupart des dispositifs - il existe actuellement des redondances dans les dispositifs dérogatoires pour un même bénéficiaire ou des différences peu justifiées entre dispositifs - et en modifiant leur barème afin d'accroître leur avantage différentiel pour les travailleurs peu qualifiés par rapport à l'allègement de charges général sur les bas salaires. En outre, éviter la superposition de plusieurs dispositifs à la fois permettrait une meilleure évaluation de leurs retombées et donc de mieux apprécier la pertinence de les prolonger ou non.

Au-delà de ces résultats sur l'efficacité des politiques ciblées sur les territoires, cette revue de la littérature a mis en évidence qu'une bonne évaluation de telles mesures devait relever plusieurs défis. Le premier est d'avoir des données géographiques adéquates. En effet, l'accès à des données géolocalisées précisément est un avantage décisif pour une bonne évaluation des programmes de zones ciblées puisque leurs contours ne correspondent pas aux frontières administratives préexistantes (municipalités, agglomération ou département). L'utilisation de données géographiques agrégées, par exemple au niveau des municipalités dans l'étude de Gobillon et al. (2012), peut générer des biais dans l'estimation. Par ailleurs, un autre inconvénient dans l'analyse de Gobillon et al. (2012) est que les données sont limitées à une seule région. Ainsi, toute évaluation de telles mesures devra mener en amont une réflexion sur la manière de définir et caractériser les zones ciblées. On doit se poser la question jusqu'à quel point de désagrégation des données doit-on aller pour avoir le moins d'erreurs de mesure dans la définition des zones ciblées.

Le deuxième défi de ces évaluations est de savoir à quoi doit on comparer les zones bénéficiant des dispositifs pour mesurer adéquatement leur efficacité. La littérature propose plusieurs stratégies pour déterminer le groupe de contrôle servant de comparaison avec les zones ciblées. Boarnet & Bogart (1996) et Hanson (2009) considèrent comme groupe de contrôle les zones éligibles aux programmes ou celles ayant candidaté mais ayant été rejetées alors que Moore (2003) et Neumark & Kolko (2010) utilisent les zones désignées comme prioritaires à des vagues ultérieures. Enfin, de nombreuses études définissent leur groupe de contrôle avec les scores de propension, soit par appariement complet [O'Keefe (2004), Busso & Kline (2008), Elvery (2009), Rathelot & Sillard (2008), Givord et al. (2012b)], soit en limitant l'échantillon à des zones dont le score appartient à des intervalles particuliers de la distribution [Bondonio & Engberg (2000), Bondonio & Greenbaum (2007), Gobillon et al. (2012)]. Là encore, au vu des données disponibles, cette question est primordiale pour la qualité de l'évaluation. Sous réserve de prendre un certain nombre de précautions, les méthodes de différences de différences avec appariement par score de propension sont séduisantes car elles sont relativement faciles à mettre en œuvre tout en tenant compte des biais de sélection et en permettant l'étude de l'hétérogénéité des effets du traitement et des effets de bouclage. Toutefois, des tests de robustesse de leurs hypothèses sont alors indispensables et doivent vraiment se faire au cas par cas.

Références

- ACOSS (2011), 'En 2010, les exonérations se stabilisent', *Bilan ACOSSTAT* **138**.
- Amghar, Y.-G. & Laloue, F. (2010), Les dispositifs dérogatoires en matière de prélèvements sociaux, Technical report, Rapport du Conseil des Prélèvements Obligatoires.
- Baudin, N., Legendre, F. & L'Horty, Y. (2010), 'Les baisses de cotisations sociales ultramarines', *Revue française d'Economie* **24**(3), 167–191.
- Baudin, N., Legendre, F. & L'Horty, Y. (2011), 'Réformer les baisses de cotisations ultramarines?', *Travail et Emploi* **1**(125), 41–52.
- Behaghel, L., Lorenceau, A. & Quantin, S. (2012), 'Tax exemptions and rural development : Evidence from a quasi-experiment', *PSE Working Papers* **2012-34**.
- Billings, S. (2009), 'Do Enterprise Zones Work ? An Analysis at the Borders', *Public Finance Review* **37**(1), 68–93.
- Boarnet, M. & Bogart, W. (1996), 'Enterprise Zones and Employment : Evidence from New Jersey', *Journal of Urban Economics* **40**(2), 198–215.
- Bondonio, D. & Engberg, J. (2000), 'Enterprise zones and local employment : evidence from the states' programs', *Regional Science and Urban Economics* **30**, 519–549.
- Bondonio, D. & Greenbaum, R. T. (2007), 'Do local tax incentives affect economic growth ? What mean impacts miss in the analysis of enterprise zone policies', *Regional Science and Urban Economics* **37**, 121–136.
- Bostic, R. & Prohofsky, A. (2006), 'Enterprise Zones and Individual Welfare : A Case Study of California', *Journal of Regional Science* **46**(2), 175–203.
- Briant, A., Lafourcade, M. & Schmutz, B. (2012), 'Can Tax Breaks Beat Geography ? Lessons from the French Enterprise Zone Experience', *PSE Working Papers* **2012-22**.
- Busso, M. & Kline, P. (2008), 'Do Local Economic Development Programs Work ? Evidence from the Federal Empowerment Zone Program', *Yale Economics Department Working Paper* **36**.
- Caliendo, M. & Kopienig, S. (2008), 'Some practical guidance for the implementation of propensity score matching', *Journal of Economic Surveys* **22**(2), 31–72.
- Cameron, A. & Trivedi, P. (2005), *Microeconometrics. Methods and Applications.*, Cambridge University Press.
- Duguet, E., L'Horty, Y. & Sari, F. (2009), 'Sortir du chômage en Ile-de-France : disparités territoriales, spatial mismatch et ségrégation résidentielle', *Revue Économique* **60**(3), 979–1010.
- Elvery, J. (2009), 'The Impact of Enterprise Zones on Resident Employment : An Evaluation of the Enterprise Zone Programs of California and Florida', *Economic Development Quarterly* **23**(1), 44–59.
- Gaschet, F. & Gaussier, N. (2003), 'Ségrégation urbaine et marchés du travail au sein de l'aire urbaine bordelaise : quelle portée pour l'hypothèse de mauvais appariement spatial?', *Cahier du GRES* **14**.
- Givord, P. (2010), 'Méthodes économétriques pour l'évaluation de politiques publiques', *Document de travail de la DESE, Insee* **G 2010 / 08**.

- Givord, P., Quantin, S. & Trevien, C. (2012*a*), ‘A Long-Term Evaluation of the First Generation of the French Urban Enterprise Zones’, *Document de travail de la DESE, Insee* **G2012/01**.
- Givord, P., Rathelot, R. & Sillard, P. (2012*b*), ‘Place-Based Tax Exemptions and Displacement Effects : An Evaluation of the *Zones Franches Urbaines* Program’, *Regional Science and Urban Economics* **forthcoming**.
- Gobillon, L., Magnac, T. & Selod, H. (2012), ‘Do Unemployed Workers Benefit from Enterprise Zones? The French Experience’, *Journal of Public Economics* **96**(9-10), 881–892.
- Gobillon, L., Selod, H. & Zenou, Y. (2007), ‘The mechanisms of spatial mismatch’, *Urban Studies* **44**(12), 2401–2427.
- Hahn, J., Todd, P. & van der Klaauw, W. (2001), ‘Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design’, *Econometrica* **69**(1), 201–209.
- Ham, J., Swenson, C., Imrohoroglu, A. & Song, H. (2011), ‘Government programs can improve local labor markets : Evidence from State Enterprise Zones, Federal Empowerment Zones and Federal Enterprise Community’, *Journal of Public Economics* **95**(7-8), 779–797.
- Hanson, A. (2009), ‘Local employment, poverty, and property value effects of geographically-targeted tax incentives : An instrumental variables approach’, *Regional Science and Urban Economics* **39**(6), 721–731.
- Hanson, A. & Rohlin, S. (2012), ‘Do spatially targeted redevelopment programs spillover?’, *Regional Science and Urban Economics* **forthcoming**.
- Heckman, J. & Hotz, V. (1989), ‘Choosing among alternative nonexperimental methods for estimating the impact of social programs : The case of manpower training’, *Journal of American Statistical Association* **84**(408), 862–875.
- Heckman, J. & Smith, J. (1997), ‘Making the Most Out of Programme Evaluations and Social Experiments : Accounting for Heterogeneity in Programme Impacts’, *Review of Economic Studies* **64**(4), 487–535.
- Heckman, J., Ichimura, H., Smith, J. & Todd, P. (1998), ‘Characterizing Selection Bias Using Experimental Data’, *Econometrica* **66**(5), 1017–1098.
- Korsu, E. & Wenglenski, S. (2010), ‘Job accessibility, residential segregation and risk of long-term unemployment in Paris Region’, *Urban Studies* **47**(11), 2279–2324.
- Lorenceanu, A. (2009), ‘L’impact d’exonérations fiscales sur la création d’établissements et l’emploi en France rurale : une approche par discontinuité de la régression’, *Economie et Statistique* **427-428**, 27–62.
- Lynch, D. & Zax, J. (2011), ‘Incidence and substitution in Enterprise Zone Programs : The case of Colorado’, *Public Finance Review* **39**(2), 226–255.
- Maguain, D. (2006), ‘Discrimination positive : un bilan des expériences américaines et européennes’, *Revue française d’Economie* **21**(2), 147–193.
- Mayer, T., Mayneris, F. & Py, L. (2012), ‘The Impact of Urban Enterprise Zones on Establishment’s Location Decisions : Evidence from French ZFUs’, *CEPR Discussion Paper Series* **9074**.

- Moore, W. (2003), 'Enterprise Zones, Firms Attraction and Retention : A Study of the California Enterprise Zone Program', *Public Finance and Management* **3**, 376–392.
- Neumark, D. & Kolko, J. (2010), 'Do enterprise zones create jobs? Evidence from California's enterprise zone program', *Journal of Urban Economics* **68**, 1–19.
- Oakley, D. & Tsao, H.-S. (2006), 'A new way of revitalizing distressed urban communities? assessing the impact of the federal Empowerment Zone Program', *Journal of Urban Affairs* **28**(5), 443–471.
- O'Keefe, S. (2004), 'Job Creation in California's enterprise zones : a comparison using a propensity score matching model', *Journal of Urban Economics* **55**, 131–150.
- Papke, L. E. (1994), 'Tax policy and urban development : Evidence from the Indiana enterprise zone program', *Journal of Public Economics* **54**, 37–49.
- Rathelot, R. & Sillard, P. (2008), 'Zones Franches Urbaines : quels effets sur l'emploi salarié et les créations d'établissement?', *Economie et Statistique* **415-416**, 81–96.
- Riedinger, N. (2009), 'Une évaluation bienvenue et rigoureuse mais dont les résultats restent à préciser', *Economie et Statistique* **427-428**.
- Rosenbaum, P. & Rubin, D. (1983), 'The central role of the propensity score in observational studies for causal effects', *Biometrika* **70**, 41–55.
- Rosenbaum, P. & Rubin, D. (1984), 'Reducing the Bias in Observational Studies Using Subclassification on the Propensity Score', *Journal of the American Statistical Association* **79**(516-524).